

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 15, 2002, 14:06:06 ; Search time 17.59 Seconds
(without alignments)
1001.557 Million cell updates/sec

Title: US-09-464-099A-70
Perfect score: 2288
Sequence: 1 MLHGASSRPATARKSSGLSG.....FMDLMAGLGAKIELSDTKAA 455

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues
al number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2282	99.7	455	1 AROA_AGRSP	Q9r4e4 agrobacteri
2	1900.5	83.1	449	1 AROA_PSES2	P56952 pseudomonas
3	884	38.6	431	1 AROA_BACHD	Q9kca6 bacillus ha
4	867.5	37.9	447	1 AROA_SVNY3	Q59975 synecocyst
5	806.5	35.2	430	1 AROA_LACLA	Q9ceu0 lactococcus
6	806	35.2	443	1 AROA_BACNO	Q46550 bacteroides
7	803.5	35.1	427	1 AROA_STRPN	Q9s400 streptococc
8	802.5	35.1	430	1 AROA_LACLC	P43905 lactococcus
9	785	34.3	428	1 AROA_BACSU	P20691 bacillus su
10	766.5	33.5	431	1 AROA_AQUAE	Q67494 aquifex aeo
11	695	30.4	410	1 AROA_THEMEA	Q9wy10 thermotoga
12	590	25.8	428	1 AROA_CAMJE	P52312 campylobact
13	577	25.2	430	1 AROA_STAAD	Q05615 staphylococ
14	521.5	22.8	429	1 AROA_HELPJ	Q9zkf7 helicobacte
15	518.5	22.7	429	1 AROA_HELPY	P56197 helicobacte
16	496	21.7	419	1 AROA_METTH	Q26860 methanobact
17	459.5	20.1	429	1 AROA_METJA	Q57925 methanococc
18	408	17.8	439	1 AROA_DEIRA	Q9rvd3 deinococcus
19	403	17.6	416	1 AROA_ARCFU	O28775 archaeoglob
20	402.5	17.6	427	1 AROA_AERPE	Q9yek9 aeropyrum p
21	402.5	17.6	427	1 AROA_YEREN	P19688 versinia en
22	395.5	17.3	428	1 AROA_YERPE	Q60112 versinia pe
23	387.5	16.9	432	1 AROA_HAESO	P52310 haemophilus
24	378	16.5	427	1 AROA_ECOLI	P07638 escherichia
25	376	16.4	427	1 AROA_SHISO	Q9zif7 shigella so
26	373.5	16.3	410	1 AROA_PYRAB	Q9v1h1 pyrococcus
27	373	16.3	432	1 AROA_PASHA	P54220 pasteurella
28	372	16.3	427	1 AROA_KLEPN	P24497 klebsiella
29	372	16.3	440	1 AROA_PASMU	Q04570 pasteurella
30	371	16.2	427	1 AROA_SHIDI	O87006 shigella dy
31	365.5	16.0	463	1 AROA_BURPS	P39915 burkholderi
32	365	16.0	428	1 AROA_EDWIC	Q9x4h2 edwardsiell
33	363.5	15.9	432	1 AROA_HAEIN	Q03421 haemophilus

34	357	15.6	427	1 AROA_SALTY	P07637 salmonella
35	348	15.2	427	1 AROA_SALGL	P22299 salmonella
36	347	15.2	427	1 AROA_SALTI	P19786 salmonella
37	323.5	14.1	516	1 AROA_PETHY	P11043 petunia hyb
38	321.5	14.1	1588	1 ARO1_YEAST	P08566 s pentafunc
39	317	13.9	441	1 AROA_CHLMU	Q9pk28 chlamydia m
40	315.5	13.8	518	1 ARO1_TOBAC	P23981 nicotiana t
41	314.5	13.7	440	1 AROA_CHLTR	O84371 chlamydia t
42	313	13.7	428	1 AROA_BUCAP	Q59178 buchnera ap
43	310.5	13.6	520	1 AROA_LYCES	P10748 lycopersico
44	308.5	13.5	1573	1 ARO1_SCHPO	Q9p7r0 s pentafunc
45	306.5	13.4	445	1 AROA_CHLPN	Q9z6m0 chlamydia p

ALIGNMENTS

RESULT 1
AROA_AGRSP STANDARD; PRT; 455 AA.
AC Q9R4E4; 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GN AROA.
OS Agrobacterium sp. (strain CP4).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Agrobacterium.
OX NCBI_TaxID:361;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-28; 47-61 AND 321-333.
RA Barry G.F., Kishore G.M., Padgett S.R., Stallings W.C.;
RT "Glyosphate-tolerant 5-enolpyruvylshikimate-3-phosphate synthases.";
RL Patent number US5633435, 27-MAY-1997.
RN [2]
RP SEQUENCE OF 1-15.
RX MEDLINE=96182485; PubMed=8598558;
RA Harrison L.A., Bailey M.R., Naylor M.W., Ream J.B., Hammond B.G.,
RA Nida D.L., Burnette B.L., Nickson T.E., Mitsky T.A., Taylor M.L.,
RA Fuchs R.L., Padgett S.R.;
RT "The expressed protein in glyphosate-tolerant soybean, 5-enolpyruvylshikimate-3-phosphate synthase from Agrobacterium sp. strain CP4, is rapidly digested in vitro and is not toxic to acutely gavaged mice";
RT J. Nutr. 126:728-740(1996).
RL -I- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate - phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC -I- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC -I- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -I- BIOTECHNOLOGY: Introduced by genetic manipulation and expressed in glyphosate-tolerant soybean, canola, cotton and maize by Monsanto. Developed to provide new weed-control options for farmers.
CC Expression of this protein in plants imparts high levels of glyphosate tolerance.
CC -I- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
DR InterPro: IPR001986; EPSP_synthase.
DR Pfam: PF00275; EPSP_synthase; 1.
DR ProDom: PD001867; EPSP_synthase; 1.
DR PROSITE: PS00104; EPSP_SYNTHASE_1; 1.
DR PROSITE: PS00885; EPSP_SYNTHASE_2; 1.
KW Aromatic amino acid biosynthesis; Transferase; Herbicide resistance; Genetically modified food.
FT CONFLICT 2 2 S -> L (IN REF. 2).
SQ SEQUENCE 455 AA; 47588 MW; 236580D08D6EF422 CRC64;

Query Match 99.7%; Score 2282; DB 1; Length 455;
Best Local Similarity 99.8%; Pred. No. 1.le-134;
Matches 454; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMEGGLASGETRITGLLEGEDVINTG 60
Db 1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMEGGLASGETRITGLLEGEDVINTG 60
QY 61 KAMQAMGARIRKEGDTWIIDGVNGGLLAPAPLDGFGNAATGCRLTWGLVGVYDFDSTFI 120
Db 61 KAMQAMGARIRKEGDTWIIDGVNGGLLAPAPLDGFGNAATGCRLTWGLVGVYDFDSTFI 120
QY 121 GDASLTKRPMGRVNLNPLREMGVQVKSDEGDRPLVTLRGPKTPTTITVRVPMASAOQVKS 180
Db 121 GDASLTKRPMGRVNLNPLREMGVQVKSDEGDRPLVTLRGPKTPTTITVRVPMASAOQVKS 180
QY 181 LLAGLNTPGITVTIEPIMTRDHTKMLQGFAGNLTVETDADGVRTIRLEGRKLTGOVID 240
Db 181 LLAGLNTPGITVTIEPIMTRDHTKMLQGFAGNLTVETDADGVRTIRLEGRKLTGOVID 240
QY 241 VPGDPSSTAFPLVAALLVPGSDVTILNLMNPTRTGLTLTLOEMGADIEVINPRLAGGED 300
Db 241 VPGDPSSTAFPLVAALLVPGSDVTILNLMNPTRTGLTLTLOEMGADIEVINPRLAGGED 300
QY 301 VADLRVRSSTLKGVTVPEDRAPSMIDEPYILAVAAAFAGATVNMGLEELRVKESDRLSA 360
Db 301 VADLRVRSSTLKGVTVPEDRAPSMIDEPYILAVAAAFAGATVNMGLEELRVKESDRLSA 360
QY 361 VANGKLNKGVDCDEGETSLVVRGRPDGKGLGNASGAATVATHLDRHRIAMSFVLMGLVSENP 420
Db 361 VANGKLNKGVDCDEGETSLVVRGRPDGKGLGNASGAATVATHLDRHRIAMSFVLMGLVSENP 420
QY 421 VTVDATMIATSFPEFMDLMAGLAKIELSDTKAA 455
Db 421 VTVDATMIATSFPEFMDLMAGLAKIELSDTKAA 455

```

Query Match 83.1%; Score 1900.5; DB 1; Length 449;
 Best Local Similarity 82.9%; Pred. No. 4.8e-111;
 Matches 373; Conservative 32; Mismatches 42; Indels 3; Gaps 1;

```

QY 1 MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMEGGLASGETRITGLLEGEDVINTG 60
Db 1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMEGGLASGETRITGLLEGEDVINTG 60
QY 61 KAMQAMGARIRKEGDTWIIDGVNGGLLAPAPLDGFGNAATGCRLTWGLVGVYDFDSTFI 120
Db 61 KAMQAMGARIRKEGDTWIIDGVNGGLLAPAPLDGFGNAATGCRLTWGLVGVYDFDSTFI 120
QY 121 GDASLTKRPMGRVNLNPLREMGVQVKSDEGDRPLVTLRGPKTPTTITVRVPMASAOQVKS 180
Db 121 GDASLTKRPMGRVNLNPLREMGVQVKSDEGDRPLVTLRGPKTPTTITVRVPMASAOQVKS 180
QY 181 LLAGLNTPGITVTIEPIMTRDHTKMLQGFAGNLTVETDADGVRTIRLEGRKLTGOVID 240
Db 181 LLAGLNTPGITVTIEPIMTRDHTKMLQGFAGNLTVETDADGVRTIRLEGRKLTGOVID 240
QY 241 VPGDPSSTAFPLVAALLVPGSDVTILNLMNPTRTGLTLTLOEMGADIEVINPRLAGGED 300
Db 241 VPGDPSSTAFPLVAALLVPGSDVTILNLMNPTRTGLTLTLOEMGADIEVINPRLAGGED 300
QY 301 VADLRVRSSTLKGVTVPEDRAPSMIDEPYILAVAAAFAGATVNMGLEELRVKESDRLSA 360
Db 301 VADLRVRSSTLKGVTVPEDRAPSMIDEPYILAVAAAFAGATVNMGLEELRVKESDRLSA 360
QY 361 VANGKLNKGVDCDEGETSLVVRGRPDGKGLGNASGAATVATHLDRHRIAMSFVLMGLVSENP 420
Db 361 VANGKLNKGVDCDEGETSLVVRGRPDGKGLGNASGAATVATHLDRHRIAMSFVLMGLVSENP 420
QY 421 VTVDATMIATSFPEFMDLMAGLAKIELS 450
Db 418 VTVDATMIATSFPEFMDLMAGLAKIELS 447

```

RESULT 3

```

ID AROA_BACHD STANDARD; PRT; 431 AA.
AC O9KCA6;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
DE enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GN AROA OR AROE OR BH1667.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
CC -!- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
CC phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN THE
CC BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC -!- SUBUNIT: Monomer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```


QY	254	RALLYPGSDVITLVNMPNTRTGLTLLTQEMGADIEVNPLRAGED----	VAOLRVRS	309
DB	242	AALVENSESLILENVGNETRIGIIEVQAMGGQLEILE-----	QDNVAKAATLVKVAS	295
QY	310	TLKGVTVPEDRAPSIMIDEYPIILAVAAFAEGATVNVNGLEELRVKESDRLSAVANGLKNG		369
DB	296	QLKGTETSGDLIPRIDEPLIALLATQAGQOTIIRDAELKVKETRIAVVANAISMG		355
QY	370	VDCDEGETSLVVRGPRDGKGLGNASGAATVATLHDIRAMSEFVLMLVGEN-PVTVDATM		428
DB	356	AKIQPTDDGMIIOG---GPKL-HAPENSINTLGDHRIGMMAIAALLVKNGEIELEAFA		411
QY	429	IATSPPEFMDLMAIGLAKI	447	
DB	412	IQTSPSPFEDDLEQLSENI	430	
RESULT	6			
ID	ARO_A	BACNO	STANDARD;	PRT; 443 AA.
AC	Q46550;			
DT	30-MAY-2000	(Rel. 39, Created)		
DT	16-OCT-2001	(Rel. 40, Last sequence update)		
DT	16-OCT-2001	(Rel. 40, Last annotation update)		
DE	3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-			
DE	enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).			
GN	ARO_A			
OS	Bacteroides nodosus (Dichelobacter nodosus).			
OC	Bacteria; Proteobacteria; gamma subdivision; Cardiobacteriaceae;			
OC	Dichelobacter.			
OC	NCBI_TaxID=870;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=VCS1001.			
RX	MEDLINE=94320795; PubMed=8045432;			
RA	Alm R.A.; Bairympie B.P.; Mattick J.S.;			
RT	"Sequencing and expression of the araA gene from Dichelobacter			
RT	nodosus."			
RL	Gene 145:97-101(1994).			
CC	-1- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =			
CC	phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.			
CC	-1- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN			
CC	THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).			
CC	-1- SUBUNIT: MONOMER (BY SIMILARITY).			
CC	-1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).			
CC	-1- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announcement/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
CC	EMBL; Z29339; CAA82544.1; -			
DR	InterPro; IPR001986; EPSP_synase.			
DR	Pfam; PF00275; EPSP_synase; 1.			
DR	ProDom; PD001867; EPSP_synase; 1.			
DR	PROSITE; PS00104; EPSP_SYNTHASE_1; 1.			
DR	PROSITE; PS00885; EPSP_SYNTHASE_2; 1.			
KW	Aromatic amino acid biosynthesis; Transferase.			
SQ	SEQUENCE 443 AA; 47444 MW; BE2243277ADEFB5 CRC64;			
Query Match	35.2%;	Score 806;	DB 1; Length 443;	
Best Local Similarity	42.2%;	Pred. No. 3.6e-43;		
Matches 185;	Conservative	74; Mismatches 159;	Indels 10; Gaps 3;	
QY	15	SSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAGAIRKEG	74	
DB	7	SQGLKRLKLVPGDKSISHRSIFMGSIAGKTKVITDILRGEDVLSITAEFRAMGVIEDKG	66	
QY	75	DTWIDGVNGGLLAEAPLDGNAATGRLTMGLVGVVDFDSTFTGDASLTKRPMGRVL	134	
DB	67	EVTVHKGKSGELKAPEKALDMGNSGTSLRLSGIAGLPFFETLFGDDSLSKRPMRVA	126	
QY	135	NPLREMGVQVQKSE-DGDRLPVTLRGPKTPPTIYRVPMSAQVKSALLAGLNTPGITTV	193	
DB	127	TPQLMGAETIGQTDKVLPMITIKGSTHLKAIDVLPVASAQVKSAVIFAALQAEGLTKV	186	
QY	194	IEPTMRDHTPEKMLQGFAGNLVETDADGVTRTLRGRKLLGQVIDVDPDSSSTAFPLV	253	
DB	187	VEKEKTRSHTEMLVQFGGELKYS-----KTLTVPGGQKLVGQKVVVPGDLSAFAFLV	241	

```

QY 76 TWIIDGVNGGLLAPDLFGNAATGCRITMCLVGVYDFDSTFIGDASLTTRPMGRVLN 135
Db 72 IVTIRGVGFLGLOPKAPLNMONGSTSMRLLAGILAAQRFESVLCODESLERPMQRIIT 131
QY 136 PLEMGVOVKSEGDRLPVTLRGPKTPTITVRPMASQVKSASVALLAGLNPGITVIE 195
Db 132 PLVOMGAKIVSHSNFTAPLHIS-RLPTGIDYALPLPSAQLKSCILLAGLADGTRLHT 190
QY 196 PIWDRHTEKMLGGGANITVETDADGVRTIRLEGRKTKTGOVIDVPGDPSSSTAPFLVAA 255
Db 191 CGISRDHTEKMLPFGGALKEKKE-----QIIVTGGQKLHGCVLDIVGDLASAAFPWVAA 245
QY 256 LVPVSGDVTILVNMNPTTGLITLQEMGADIEVINPLAGGEDVADLRVRSSTLKGVT 315
Db 246 LIAPRAEVIRNVGINPTRAAITLLOKMGRIELHHQFWEAGPVADIVVHHSKURGIT 305
QY 316 VPEDRAPSMIDEPYILAVAAFAEGATVNMGLFELRVKESDRLSAVANGKLVGDCDEG 375
Db 306 VAPEWIANAIDELPIFFIAACAEGTTFVGNLSLSELRVKESDRLAAMAQNLQTLGVACDVG 365
QY 376 ETSLVVGRPGDKGLGNASGAAYAVHLDHRIAMSFYMLGVLSENPTVDDATMIATSPFE 435
Db 366 ADFIHYGRSDHQFL-----PARVNSFGDHRIAMSLAVAGVRAAGELLIDDGAVAAVSMFQ 421
QY 436 FMDLMAGLGAKEIELSDTK 453
Db 422 FRDFAAAGMNVGERDAK 439

RESULT 7
AROA_STRPN STANDARD; PRT; 427 AA.
AC Q9SA400:
DT 30-MAY-2000 (Rel. 39, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
  enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
OS AROA OR SPI371.
GN Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX STRAIN=NCIMB 40794 / 0100993;
RX MEDLINE=20069365; PubMed=10601870;
Du W., Wallis N.G., Mazzulla M.J., Chalker A.F., Zhang L., Liu W.-S.,
Kallender H., Payne D.J.;
3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
  enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).";
RT Eur. J. Biochem. 267:222-227(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=TIGR4;
RX MEDLINE=21357209; PubMed=11463916;
RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,
RA Umayam A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Holtzapple E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L.,
RA McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,
RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus
  pneumoniae.";
RL Science 293:498-506(2001).
CC -!- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
  phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
  THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).

```

```

CC -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; AF169483; AAD45819.1; -.
DR TIGR; SP1371; -.
DR InterPro; IPR001986; EPSP_synthase.
DR Pfam; PF00275; EPSP_synthase; 1.
DR ProDom; PD001867; EPSP_synthase; 1.
DR PROSITE; PS00104; EPSP_SYNTHASE_1; 1.
DR PROSITE; PS00885; EPSP_SYNTHASE_2; 1.
KW Aromatic amino acid biosynthesis; Transferase; Complete proteome.
FT CONFLICT 13 13 S -> I (IN REF. 1).
FT CONFLICT 71 71 I -> V (IN REF. 1).
FT CONFLICT 201 201 K -> Q (IN REF. 1).
FT CONFLICT 303 303 G -> C (IN REF. 1).
SQ SEQUENCE 427 AA; 45766 MW; 45CE6F4D0D1C7B70 CRC64;

Query Match 35.1%; Score 803.5; DB 1; Length 427;
Best Local Similarity 42.8%; Pred. No. 4.9e-43;
Matches 184; Conservative 72; Mismatches 155; Indels 19; Gaps 7;

QY 18 LSGTVRIPGDKSIHRSFMFGGLASGETRITGLLEGEDVINTKAMQAMGARIKSGDTW 77
Db 10 LHGSIRVPGDKSIHRSIIIFGSLAEGTKVYDILRGEDVLSTMQVFRDLGVEIEDKGVI 69
QY 78 IDGVGVNGGLLAPDLFGNAATGCRITMCLVGVYDFDSTFIGDASLTTRPMGRVLNPL 137
Db 70 TIQGVGMAGLKAPQNALNMNGNSCTIRLISGLVLAGADFEMFGDDSLSKRPMDRVTLPL 129
QY 138 REMGVQVKS-EGDRLPVTLRGPKTPTITVRPMASQVKSASVALLAGLNPGITVIEP 196
Db 130 KMGVSISSQTERDLPRLKTKNLRPIHYELPIASQVKSALMFAALQAKGESVIEIK 189
QY 197 IFRDHTKMLQGFANLTVETDADGVRTIRLEGRKTKTGOVIDVPGDPSSSTAPFLVAAAL 256
Db 190 EYTRNTEEDMLKQFGHLSV----DG-KKITVQGPQKLTGQKVVVPGDISSAFLVAGL 244
QY 257 LVPVSGDVTILVNMNPTTGLITLQEMGADIEV--INPRLAGEDVADLRVRSSTLKG 314
Db 245 IAPNSRLVLQNVGINETRTGIIDVIRAMGKLEITEIDPVAKS---ATLIVESDGLKT 300
QY 315 TVPEDRAPSMIDEPYILAVAAFAEGATVNMGLFELRVKESDRLSAVANGKLVGDCDE 374
Db 301 EIGGALIPRLIDELPIALLATQAGVTVIKDAEELKVKETDRIQVYVADALNSMGADITP 360
QY 375 GETSLVVGRPGDKGLGNASGAAYAVHLDHRIAMSFYMLGVLSENPTVDDATMIATSF 433
Db 361 TADGMIKKGK-----SALHGARVNTFGDHRIGCMTIAALLVADGEVELDRAEINTSY 414
QY 434 PEFMDLMAGL 443
Db 415 PSFFDDLES 424

RESULT 8
AROA_LACLC STANDARD; PRT; 430 AA.
ID AROA_LACLC
AC P43905;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
  enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GN AROA.

```

Lactococcus lactis (subsp. cremoris) (Streptococcus cremoris).
 CC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 CC Lactococcus.
 CC NCBI_TaxID=1359;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=MG1363 / F15876;
 RX MEDLINE=95124293; PubMed=7823907;
 RC Griffin H.G., Gasson M.J.;
 RA "Genetic aspects of aromatic amino acid biosynthesis in Lactococcus
 RT lactis.";
 RL Mol. Gen. Genet. 246:119-127(1995).
 CC -1- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
 CC phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
 CC -1- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHOLISMATE WITHIN
 CC THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
 CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -1- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X78413; CAA53180.1; .
 CC InterPro: IPR001986; EPSP_synthase.
 CC Pfam: PF00275; EPSP_synthase; 1.
 CC ProDom: PD001867; EPSP_synthase; 1.
 CC DR PROSITE: PS00104; EPSP_SYNTHASE_1; 1.
 CC DR PROSITE: PS00883; EPSP_SYNTHASE_2; 1.
 CC KW Aromatic amino acid biosynthesis; Transferase.
 CC Aromatic amino acid biosynthesis; Transferase.
 CC SEQUENCE 430 AA; 45804 MW; C5A197A49072C9D5 CRC64;
 CC -----

Query Match 35.1%; Score 802.5; DB 1; Length 430;
 Best Local Similarity 41.2%; Pred. No. 5.7e-43;
 Matches 181; Conservative 79; Mismatches 158; Indels 21; Gaps 7;

Qy 15 SSGLSGTVRIPGDKSISHRSFPMFGGLASGETRITGLLEGEDVINTGKAMQAGAIRKEG 74
 Db 7 SOGLKGRLLKVPDGSISHSIFMSIFGSIAGKTKTIHDLRGEDVLSITIEAFRALGVEIEDDG 66
 Qy 75 DTWIIDGVNGGLLAPEALPDGNAATGCRLTMLGVGVDFDSTFIGDASLTKRPMGRVL 134
 Db 67 QVITVHGQGISKLKEPEKALDMGNSGTTSLRSLGILAGLPPEATLFGDSSLKRPMDRVA 126
 Qy 135 NPLREMGVQVKSE-DGDRLLPVLTRGPKTPPTTYRVPMSAQVKSALLAGLNTPGITTV 193
 Db 127 TPLQMMGAIEVQOTDKVKLPMTIKGSAHLKAIDYILPVASAQVKSAVIFAALQAEGLTKV 186
 Qy 194 IEPIMTRDHTERKMGOGFANLTVDADGVTRIRLEGRKLTGQVIDPDPSSSAFPLV 253
 Db 187 VEKEXTRSHREMLVQFGGEITVSD-----KTIPLPGGQKLGGQEVTPGDISSAAFWLV 241
 Qy 254 AALLVPGSDVTILNVLNMPTRGLTILTQEMGADIEVINPRLAGGED----VADLRVRS 309
 Db 242 AGLVENSGLIILENGINETRIGILEVIOAMGGQLEILE-----QDEVAKAATLKVKAS 295
 Qy 310 TLKGVTVPEDRAPSMIDEYPIILAAVAFAGATVMNGLEELRVKESDRLSAVANGLKNG 369
 Db 296 QLKGTIEISGDLPIRLIDELPIIALLATQAEKGTIIRDAEALVKVETDRIAVVADALNSMG 355
 Qy 370 VDCDEGETSLVVRGPDGKGLGNASCAAVATLHDRIAMSFLVMGLVSEN-PVTVDATM 428
 Db 356 ANIEPTDGMGIQGG---GTYKL-HAPENALITLGDHRIAGMVAIALLVENGETELERAEE 411
 Qy 429 IATSFPEFMDLMAGLGAKI 447
 Db 412 IOTSYPSFFDLEKLSGNL 430

```
Db 239 AAGAMPNSRIVLKNVGLNPTRTGDIIDVLONMGAKLEIKPSADSAGPEYGDLLIETSSLK 298
Qy 313 GVTVPEDRAPSIDYEPILAVAAFAEAGATVNGLEELRVKESDRLSAVANGKLNQVDC 372
Db 299 AVEIGGDIPIRLIDIPILATIAATQAGTTVTKDAELKVKTETNRIDTVSELRLGABEI 358
Qy 373 DSGETSLVVRGPRDCKGLGNAGAAVATHLDRHIAVMSFLVNGLVSENPVTVDDATMIATS 432
Db 359 EPTADGMKYVQKTUQG-----GAAVSSHGDRHGMGLGIASCITEEPIETHTDAIHVS 413
Qy 433 FPEFMDLMLAGLAK 446
Db 414 YPTFEHLNLSKK 427

RESULT 10
ARO_AQUAE STANDARD; PRT; 431 AA.
O67494;
30-MAY-2000 (Rel. 39, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GN AROA AQ 1536.
OS Aquifex aeolicus.
OC Bacteria; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
[1]
SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
Graham D.E., Overbeek R., Sneed M.A., Keller M., Ruvaj M., Huber R.,
Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
aeolicus."
RL Nature 392:353-358(1998).
CC -!- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL: AE000744; AAC07443.1; -
DR InterPro: IPR001986; EPSP_synthase.
DR Pfam: PF00275; EPSP_synthase; 1.
DR ProDom: PD001867; EPSP_synthase; 1.
DR PROSITE: PS00104; EPSP_SYNTHASE_1; 1.
DR PROSITE: PS00885; EPSP_SYNTHASE_2; 1.
KW Aromatic amino acid biosynthesis; Transferase; Complete proteome.
SQ SEQUENCE 431 AA; 47793 MW; EF842512EBE41D2A CRC64;
```

Query Match 33.5%; Score 766.5; DB 1; Length 431;
Best Local Similarity 40.7%; Pred. No. 9,7e-41;
Matches 174; Conservative 73; Mismatches 169; Indels 11; Gaps 6;

```
Qy 14 KSSGLSGTVRIPEDKSIHRSFNGGLASGETRTGCLLEGEDEVINFGKAMQAGARIRKE 73
Db 6 KIKRVKGLRVPDSKSIHTRAPILGALASGETLVKRPKLTSGDTLATLELTKAIRTKVREG 65
```

```
Qy 74 GDTWIIDGVNGGLLAPLEAPLDFGNAATGCRLTMTGLGVYVDFDSTFIGDASLTKKRPMGRV 133
Db 66 KEEVLIEG-RNTTELEPHDVLDKAKSGTTARTIMSGVGLSTQPFESVLTGDESUKRPMRLRV 124
Qy 134 LNPLREMGVQVKS-EDGDRLPVTLRGKPTPTTITRVPMASQAQVKSAVILLAGLINTPGITT 192
Db 125 VEPLREMGAKIDGREGNKLPITAIRGNL-KGISYFNKKSSQAQVKSALLLAGLRABGMTE 183
Qy 193 VTEPTMTDRHTKMLQGFQGANLTVETDADGVTRTIRLEGRGKLTGGQVIDVPGDPSSSTAPPL 252
Db 184 VVEPYLSRDHTERMLKLFQAEVITTIPEERG-HIVKIKGQQLQGTVEYCPADPSSAAYFA 242
Qy 253 VAALLVPGSDVTILNVMNPRTGLTLTQEMGADIEVINPLAGGEDVADLRVRS-STL 311
Db 243 ALATLAPEGEIRLKEVLLNPTPDGFRKLIEMGGDISFENYRELSNEPNADLVVRVDNL 302
Qy 312 KGVTPEDRAPSIDYEPILAVAAFAEAGATVNGLEELRVKESDRLSAVANGKLNQVDC 371
Db 303 KPVKVSPEEVPITLDEIPILAVMAFADGVSEVKGAKELRYKESDRKAIVTNLRKLGVO 362
Qy 372 CDEGETSLVVRGPRDCKGLGNAGAAVATHLDRHIAVMSFLVNGLVSENPVTVDDATMIAT 431
Db 363 VEEFEDGFAIHGTKEIK-----GGVIETFKDHRITAMAFVLGLVVEEVIIDHPCEVTV 416
Qy 432 SPPEFMD 438
Db 417 SYPERWE 423

RESULT 11
ARO_AQUAE STANDARD; PRT; 410 AA.
AC O9WYIO;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GN AROA OR TW0345.
OS Thermotoga maritima.
OC Bacteria; Thermotogales; Thermotoga.
OX NCBI_TaxID=23336;
[1]
SEQUENCE FROM N.A.
RC STRAIN=MS98 / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
Haft D.H., Hickley E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
Saizberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
genome sequence of Thermotoga maritima."
RL Nature 399:323-329(1999).
CC -!- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL: AE001715; AAD35431.1; -
DR TIGR: TM0345; -
```

Thu Aug 15 13:49:59 2002

[illegible]

```

RESULT      12
AFROA_CAMJE STANDARD;          PRF;    428 AA.
ID AFROA_CAMJE
AC P52312; Q9PP36;
DT 01-OCT-1996 (Rel. 34, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
FT 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5'-
GE enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GN AROA OR CJO895C.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
OC Campylobacter.
OX NCBI_TaxID=197;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=81116;
RX MEDLINE=971128776; PubMed=8973316;
RA WOESTEN M.M.S.M., DUBBINK V.H.J., van der ZELIJST B.A.M.;
RT "The aroA gene of Campylobacter jejuni.";
RL Gene 181:109-112(1996).
[2]
RN SEQUENCE FROM N.A.
RC STRAIN=NCTC 11168;
RX MEDLINE=20150912; PubMed=10688204;
RA Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,
RA Jagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
RA Whitehead S., Barrell B.G.;
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni

```

RT reveals hypervariable sequences. ",
RL Nature 403:665-668(2000).
CC -1- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate -
CC phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
CC THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X893771; CAA61554.1; ..
CC EMBL; ALJ39076; CAB73153.1; ..
DR InterPro; IPR001986; EPSP_synthase.
DR Pfam; PF00275; EPSP_synthase; 1.
DR ProDom; PD001867; EPSP_synthase; 1.
DR PROSITE; PS0104; EPSP_SYNTHASE_1; 1.
DR PROSITE; PS00885; EPSP_SYNTHASE_2; 1.
CC AROMATIC amino acid biosynthesis; Transferase; Complete proteome.
KW CONFLICT 152 152 K -> N (IN REF. 1).
FT CONFLICT 158 158 F -> Y (IN REF. 1).
FT CONFLICT 179 179 T -> N (IN REF. 1).
FT CONFLICT 183 183 T -> A (IN REF. 1).
FT CONFLICT 217 217 N -> S (IN REF. 1).
FT CONFLICT 241 241 V -> A (IN REF. 1).
CC SEQUENCE 428 AA; 47288 MW; 465E2B50F39FC5AB CRC64;

Query Match 25.8%; Score 590; DB 1; Length 428;
best local similarity 34.0%; pred. NO. 8.5e-30;

[illegible]

RESULT 13


```

AROA_STAAU STANDARD; PRT; 430 AA.
ID AC Q05615;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
DE enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GN AROA.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Staphylococcus.
OX NCBI_TaxID=1280;
[1]
SEQUENCE FROM N.A.
RC STRAIN=NCCT 8325-4;
RC MEDLINE=93381456; PubMed=8371108;
RA O'Connell C.M., Pattee P., Foster T.J.;
RT "Sequence and mapping of the aroA gene of Staphylococcus aureus
8325-4.";
J. Gen. Microbiol. 139:1449-1460(1993).
CC -1- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
CC phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC -1- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
CC THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
EMBL; L05004; AAA71897.1; -
DR InterPro: IPR001986; EPSP_synthase.
DR Pfam: PF00275; EPSP_synthase; 1.
DR ProDom: PD001867; EPSP_synthase; 1.
DR PROSITE: PS00104; EPSP_SYNTHASE_1; 1.
DR PROSITE: PS00895; EPSP_SYNTHASE_2; 1.
KW Aromatic amino acid biosynthesis; Transferase.
SQ SEQUENCE 430 AA; 46852 MW; E5FC878EALC23C20 CRC64;
-----
Query Match 25.2%; Score 577; DB 1; Length 430;
1st Local Similarity 33.8%; Pred. No. 5.5e-29;
Matches 144; Conservative 82; Mismatches 179; Indels 24; Gaps 8;
QY 15 SSGLSGTVRIPGKSIHSRSMFGGLASGSTRITGLLEGEDVINTGKAMQAMGARIRKEG 74
DB | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 10 SGPLKGEIEVPGDKSMTHRAIMLASLAEGVSTYTKPLLGEDCRTMDFRHLGVEIKEDD 69
QY 75 DTWIDGVNGGLIAPAPLDFGNAAATGCRLTGMLGVGYDFDSTFIGDASLTRKPMGRVL 134
DB | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 70 EKLVTSPGY-QVNTPHQVLYTNGSGTTRLLAGLLSGLGNESVLGSDVSTIGRKPMDRVL 128
QY 135 NPLREMGVQVKSGDGRPLVTLRGPKTPPTITYRVPMAQVKSVAVLLAGLNPITGTTVI 194
DB | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 129 RPLKMDANTIEGIEDNTYPLIIK-PSVIKGINQYQMEVASAQVKSAILFASLFKEPTIIK 187
QY 195 EPIMTDRTHKMLQGF-----GANLTVETDADGVRTIRLEGRGLTKQVIDVCGDPSTA 249
DB | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 188 ELQVSRNHTETMFKHFNPIETAEGLSINTTPEAIRIKPAD-----FHPVGDITSSAA 239
QY 250 FPLVAALLVPGSDVTIINVLNMTPTRTGLITLQBMGADIEVINPRLAGGEDVADLRV-S 308
DB | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 240 FFIIVAALITPGSDVTIHNVGINTRSGIIDIIVKMGNIQLFN-QTTGAETASIRIQYT 298
QY 309 STLKGVTVPEDRAPSMTDYEPIIAVAFAAEAGATVMNGLLERLKRVKSIVANGKLKN 368
DB | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

```

Db 299 PMLQPTITIEGELVPKAIDELPAVIALLCQAVGTSTIKDAEELKVKENRIRDTTADMNL 358
 QY 369 GVDCEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSLFVLGVSENPVTVDATM 428
 Db 359 GFELQPTNDGLIHH--PSEKTNATDILT---DHRIGMMLAVACVLSSEPVKIKQFDA 411
 QY 429 IATSFPEFM 437
 Db 412 VNVSPGFL 420
 RESULT 14
 AROA_HELPJ
 ID AROA_HELPJ STANDARD; PRT; 429 AA.
 AC Q9ZKF7;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
 DE enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
 GN AROA OR JHP0980.
 OS Helicobacter pylori J99 (Campylobacter pylori J99).
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
 OC Helicobacter.
 OC NCBI_TaxID=85963;
 RN [1]
 RZ SEQUENCE FROM N.A.
 RX MEDLINE=99120557; PubMed=9923682;
 RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
 RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
 RA Tunman P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
 RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
 RA Trust T.J.;
 RT "Genomic sequence comparison of two unrelated isolates of the human
 RT gastric pathogen Helicobacter pylori.";
 RL Nature 397:176-180(1999).
 CC -!- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
 CC phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
 CC -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
 CC THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
 CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE001527; AAD06557.1; -
 DR InterPro: IPR001986; EPSP_synthase.
 DR Pfam: PF00275; EPSP_synthase; 1.
 DR ProDom: PD001867; EPSP_synthase; 1.
 DR PROSITE: PS00104; EPSP_SYNTHASE.1; 1.
 DR PROSITE: PS00885; EPSP_SYNTHASE.2; 1.
 KW Aromatic amino acid biosynthesis; Transferase; Complete proteome.
 SQ SEQUENCE 429 AA; 47167 MW; 92724C4A25752741 CRC64;
 Query Match 22.8%; Score 521.5; DB 1; Length 429;
 Best Local Similarity 33.68; Pred. No. 1.5e-25;
 Matches 145; Conservative 73; Mismatches 166; Indels 47; Gaps 13;
 QY 27 DKTSIHSRFMFGIASGETRITGLLEGGEDVINTGKAMQAMGARIRKEG-----DTWI 78
 Db 10 DKSLSHRAVIFSLAQKPCFVRNLMGEDCLSSLEIAQLGAKVENTAKNSFKITPPTTI 69
 QY 79 IDGVNGGLLIAEAPLDFGNNAATGCRITMGVLGVYDFDSTFIGDASLTGKPMGRVNLPLR 138
 Db 70 KE-----PNKILNCNNGTSMRWLYSGLLSAQKGLFVLGSDNSLNARPMKRIEPLK 120

139 EMGVQVKS-EDGDRLPVTLRGPKPTPTITVYRVPMASQVKSAVLLAGLNTPGITTVIEPI 197
121 AFGAKILGREDNHFAIVGGPLKA-CDVESPISASQVKSFAFTLSALQAQGISAYKESE 179
198 MTRDHTKMLQGFANLTVETDAGVTRIR-LEGRGKLTGQVIDVPDGPSTAPPLVAAL 256
180 LSRNHTIEMKSLGANI---NQDGVKISPLE--KPLESFDFIANDPSSAFVLTALACA 234
257 LVPQSDVTILNVLNPTRTGLILTLOEMGADIE-VINPRLAGGEDVADLRVRSSTLKGVT 315
235 ITPKSRLLKNVLLNPTRIEAFVLLKMGAEIVYQSK--DLEVIGDIYIEHAPLKAIS 292
316 VPEDRAPNIDEYPIIAVAAAFASGATVMNGLEELRVKESDRLSAVANGLKNGVDCDEG 375
293 IDQNTA-SLIDEIPALSIAMLFAGKSMVRNAKDLRAKESDRIRKAVVSNFKALGIECEF 351
376 ETSLVVRGRPDGKGLGNAS-----GAAVATHLDHRIAMFLVNGVLYSENPTVDD 425
352 EDGFI-----EGLGDASQLKQHSKTKPIIKSFNDHRIAMFAVLTAL--PLEIDN 403
426 ATMATSPPEF 436
404 LECANISFPTE 414

RESULT 15
AROA_HELPY
ID AROA_HELPY STANDARD; PRT; 429 AA.
AC P56197;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
DE enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GN AROA OR HP0401.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Kervavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT Pylori";
RL Nature 388:539-547(1997).
CC -1- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
CC phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC -1- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
CC THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
CC or send an email to license@isb-sib.ch.
CC -----
CC EMBL; AE000556; AND07470.1; "

TIGR: HP0401; .
DR InterPro: IPR001986; EPSP_synthase.
DR Pfam: PF00275; EPSP_synthase; 1.
DR ProDom: PD001867; EPSP_synthase; 1.
DR PROSITE: PS00104; EPSP_SYNTHASE_1; 1.
DR PROSITE: PS00885; EPSP_SYNTHASE_2; 1.
KW Aromatic amino acid biosynthesis; Transferase; Complete proteome.
SQ SEQUENCE 429 AA; 47240 MW; 19545753E081FDAE CRC64;

Query Match 22.7%; Score 518.5; DB 1; Length 429;
Best Local Similarity 33.5%; Pred. No. 2.3e-25;
Matches 145; Conservative 74; Mismatches 171; Indels 43; Gaps 13;

QY 27 DKSISHRSMFVGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEG-----DTWI 78
DB 10 DKSLSHRAVIFSLAQAQPCFVRNFMGDCDLSLEIAQNLGAKVNTAKNSFKITPPTTI 69
QY 79 IDGVNGGLLAPEAPLDFGNAATGCRLTGMVLGVYDFDSTFIGDASLTGRPMGRVNLPLR 138
DB 70 KE-----PNKILNCNNSGTTMRLYSGLLSAQKGLFVLSGDNSLNARPMKRIIEPLK 120
QY 139 EMGVQVKS-EDGDRLPVTLRGPKPTPTIT---YRVPMASQVKSAVLLAGLNTPGITTVI 194
DB 121 AFGAKILGREDNHFAIVILG----SPLKACHYESPIASQVKSFAFTLSALQAQGASTYK 176
QY 195 EPIMTRDHTKMLQGFANLTVETDAGVTRIR-LEGRGKLTGQVIDVPDGPSTAPPLV 253
DB 177 ESELSRNHTEIMKSLGADI---HNQDGVKISPLE--KPLEAFDFIANDPSSAFFAL 231
QY 254 AALLVPGSDVTILNVLNPTRTGLILTLOEMGADIE--VINPRLAGGEDVADLRVRSSTL 311
DB 232 ACATTPKSRLLKNVLLNPTRIEAFVLLKMGASIEVIAQSKDL---EMIGDIYVEHAPL 288
QY 312 KGVTVVPEDRAPNIDEYPIIAVAAAFASGATVMNGLEELRVKESDRLSAVANGLKNGVD 371
DB 289 KAINIDQNTA-SLIDEIPALSIAMLFAGKSMVRNAKDLRAKESDRIRKAVVSNFKALGIE 347
QY 372 CDEGETSLVVRGRPDGKGL---GNASCAAVATHLDHRIAMFLVNGVLYSENPTVDDAT 427
DB 348 CEEFEDGEYVEGLEDISPLKQRFSSRIKPLIKSFNDHRIAMFAVLTAL--PLEIDNLE 405
QY 428 MATSPPEFMDLM 440
DB 406 CANISFPQFKHLL 418

Search completed: August 15, 2002, 14:06:07
Job time: 406 sec

THIS PAGE BLANK (USPTO)

QY 361 VANGKLINGVDCDEGETSLVVRGRPDGKGLGNAGAAVATHLHDRIAMSFVYMGVLENP 420
 |||: ||||| ||| ||| ||||| ||| ||| ||||| ||||| ||| |||
 Db 391 VARGLEANGVDCTEGEMSLTVVRGPGGKGLG---GGTVGTHLHDRIAMSFVYMGVLEAKP 447
 QY 421 VTVDATMIATSPPEFMDLMAGLGAKI 447
 ||||: ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 448 VTVDSTMIATSPPEFGMMAGLGAKI 474

RESULT 2
 C87694
 3-phosphoshikimate 1-carboxyvinyltransferase [imported] - Caulobacter crescentus
 C:Species: Caulobacter crescentus
 C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
 C:Accession: C87694
 R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
 B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
 n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
 A:Title: Complete Genome Sequence of *Caulobacter crescentus*.
 :Reference number: AB7249; MUID:21173698; PMID:11259647
 :Accession: C87694
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-443 <SPTO>
 A:Cross-references: GB:AE005673; NID:g13425333; PIDN:AAK25551.1; GSPDB:GNO0148
 C:Genetics:
 A:Gene: CC3589
 A:Superfamily: 3-phosphoshikimate 1-carboxyvinyltransferase; 3-phosphoshikimate 1-carbox

B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; KoloD.
n. J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A>Title: Complete Genome Sequence of Caulobacter crescentus.
AReference number: AB7249; MUID:21173698; PMID:11259647
Accession: C87694
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-443 <STO>
A:Cross-references: GB:AE005673; NID:g13425333; PIDN:AAK25551.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC3589
C:Superfamily: 3-phosphoshikimate 1-carboxyvinyltransferase; 3-phosphoshikimate 1-carboxy

	Query Match	49.6%	Score 1135;	DB 2:	Length 443;
	Best Local Similarity	55.8%;	Pred. No. 4.4e-65;		
	Matches	251;	Conservative	45;	Mismatches 138;
				Indels	16;
				Gaps	6;
Qy	2 LHGASSRPATARKSSGLSGTGVIRPDGKSISHRSFPMGGLASGETRITGLLEGEDVINTGK	61			
Dd	3 LAGLKSAPGA-----LRGITVPADCKSISHRSMLGNALATTGTTVESLLGGDDVLATAR	57			
Qy	62 AMQANGARIKREG-DTWIIDGVNGGLLAPAPLDGFNAATGCRLTWGLVGVDYDFDSFEI	120			
Dd	58 AMQAFGARIEREGVGRWRIE--GKGGFPEPVDVDCGNAGTGVRLLMGAAAGFAMCATFT	115			
Qy	121 GDASTLKPMKRVLNPLREMGVQVKSDGRLPYTLRGPKPTPTITYRVPMAQAQVKSASV	180			
Dd	116 GDQSURGRPVMGLDPLARMGATWLGRDKRLPLTKGNL-RGLNYTLPWASAQAQVKSASV	174			
Qy	181 LLAGLNTFGITTIVIEPIIMTRDHTEKMGOGGANLTVTETDADGVRT---IRLEGRGLTQG	237			
Dd	175 LLAGLHAEGVEVIEPEATRTOHTRMLRAFAEVIYVEDRKAGDKTFRHVRLPEGOQLTGT	234			
Qy	238 VIDVPGDSPSTAFFPLVAALLVPGSDVTILNVLMPNTRTGLILTLOEMGADLEVINPRLAG	297			
Dd	235 HVAVPGDPSSAAFPLVAALLVPGSEVTVGYMNLRELRTGLFTTLOEMGADLVISNVRSAS	294			
Qy	298 GEDVADRLRVSRSLTKGVTPVEDRAPSMIDEYPILAVAAFAFEGATVMNGLEELRVKESDR	357			
Dd	295 GEEVGDIITARYSQLKVVPVPERAPSMDIYEPILAVAAFAFASGETVMRGVGMRYKESDR	354			
Qy	358 LSNAVANGLKNGVDCDEGETSVLVRGRPDGKLGNASGAATAATHLDHRIANSFLVMGLVS	417			
Dd	355 ISLTANGLKACGVVVVEPEGFIIV----TGTCPPKGGATVYTHGDHRIANSHLLIGMAA	410			
Qy	418 ENPVTDDATMTATSPPEFMFLMAGLGAKI	447			
Dd	411 OAFVAVDEPEGMTATSPFGADLMRWGLGATL	440			

```

Db      235  HVAVPGDPSAAFPILVAALIVPGSVTVVEGVMLNELRTGLFTTLOEMGADILVINSRVAS 294
Qy      298  GEDVADLRVRSSTLKGTVVPEDRAPSMIDEPYILAVAAAFAGEGATVMNGLBELRVKESDR 357
Db      295  GEEVGDIARYSQLKGVVPPERAPSMIDEPYILAVAAAFASGETVMRGVGMERVKESDR 354
Qy      358  LSAVANGKLNGVDCDEGTSVLVVRGPDGKGLGNASGAAYATHLDHRIANSFLVYGLVS 417
Db      355  ISLTANGLKACGVQVVEEPEGV---TGTGPPKGGATVVTGHDHRIANSHLILGMAA 410
Qy      418  ENPVTDDATMTATSPPEFMDLMAGLGAKI 447
Db      411  QAEVAVDEPGMIATSPFGADLMRGLGATL 440

RESULT      3
E83250
still frameshift 3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTTRANSFERASE prephenate dehydrogenase I
C:Species: Pseudomonas aeruginosa
15-Feb-2000 #acruaginosa revision 15-Sep-2000 #text change 31-Dec-2000

```

[illegible]

Query Match 37.9%; Score 867.5; DB 2; Length 447;
Best Local Similarity 48.1%; Pred. No. 5.2e-48;
Matches 203; Conservative 62; Mismatches 144; Indels 13; Gaps 7;

ay 18 LSGTVRIPGDKSISHRSFMFGGLASGETRTIIGLEGEDVINTKAMQAMGARIRK-EGDT 76
||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
bb 26 LTGRRLVRPGDKSIASHRALMLGAITGETIIIGLLIGEDPRSTAHCFRAMGAEIFSELNSEK 85
||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
ay 77 WIIDGVNGGILLAPPEALDFGNAAATCGCKLTMGLV-GVVDFDSTFIGDASLTKRPMGRVLN 135
||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
bb 86 IIVOGRLGQLQEPSTVLDAAGNSGTTMRMLGLLAGQKDCFLTFTVGDDSLRHRRPMSRVIQ 145
||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
ay 136 PLREMGVOV-KSEGDRLPVTLRGPKTPTITYRPVMAAQVKSAVILLAGLNTGPITTVI 194
||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
bb 146 PLOQMGAARWNSKGKFAPLAVQSOL-KPIHYHSPIASAQVKSCLLLAGLTTEGDTTVT 204
||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
ay 195 EPIPTRDRHTEKMLOGFGANLVETADADGVRTIRLEGKLTGOVIDVPDSSSTAFFPLVA 254
||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
bb 205 EPALSRDHSERMLQAFGAKLTID---PYTHSVTVHGPAHLTGQRVVVFDPDISAFAFWLVA 261
||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
ay 255 ALLVPGSDVTILNLNMNTRTGLTILOEMGADIENINPLAGGEDVADLRVRSSTLKGV 314
||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
bb 262 ASILPGESELLVENGINTRTCGLEVLQAOMGADIIPENERLVTGEPVADLRVRASHLQGC 321
||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
ay 315 TVPEDRAPSMIDEYPILAFAAFASGATVMNGLEELRVKESDRLSAVANGKLNGVDCE 374
||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
bb 322 TFGGEIIPRLIDEIPILAFAFAAECTTRIEDAAELRVKESDRLLAAIASELGKMGAKVTE 381
||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
ay 375 GETSLVWRGPPGKGLGNASGRAATHLDHRITAMSFVLMGLVSENPTVDDATMATISFP 434
||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
bb 382 FDDGLEIOG---GSPL---QGAEVDSLTDHRIAMALAIAALGGGOTIINRAEAAAISYP 435
||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
ay 435 EF 436
||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
db 436 EF 437

RESULT 7
AC1687
5-enolpyruvylshikimate-3-phosphate synthase homolog aroE [Imported] - Listeria innocua
C.Species: Listeria innocua
C.Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C.Accession: AC1687
R.Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker,
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkhat, G.; Madueno, E.; Maitournam, A.; Ma,
C.; Schluerter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,
K.; Title: Comparative genomics of Listeria species.
A.Reference number: AB1077; WUID:21537279; PMID:11679669
A.Accession: AC1687
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-428 <GLA>
A.Cross-references: GB:AL92022; PIDN:CAC97267.1; PID:g16414538; GSPDB:GN00178
A.Experimental source: strain Clpl1262
C.Genetics:
C:A Gene: aroE
C:Superfamily: 3-phosphoshikimate 1-carboxyvinyltransferase; 3-phosphoshikimate 1-carbo-

Query Match 35.6%; Score 815.5; DB 2; Length 428;
Best Local Similarity 40.8%; Pred. No. 1e-44;
Matches 173; Conservative 83; Mismatches 155; Indels 13; Gaps 5;

ay 17 GLSGTVRIPGDKSISHRSFMPGGIASGETRTITGLLEDGVINTKAMQAMGARIRKEGDT 76
||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
db 9 GLVEITVPDGKSIMSHRSIMFGAIEGKTIVRHFLRADDCDLGTIKAFKALGVKIETEEDDE 68
||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
ay 77 WIIDGVNGGILLAPPEALDFGNAAATGRLTMGLVGVVDFDSTFIGDASLTKRPMGRVLNP 136
||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :


```
Db 296 QLKGTGSEIDLPRIIDELPIIALATOGQTIIDHRAELKVKEDRIAVVANALNSMG 355
QY 370 VDCDEGETSLVVRGRPDGKGLGNASGAATHLDHRIAMSLVGMGLVSEN-PVTVDATM 428
Db 356 AKIQPTDDGMIIOG---GFKL-HAPENSINTGLDHRIGMAAIAALLVKNGEIELERA 411
QY 429 IATSPFPEMDLMAGLAKI 447
Db 412 IQTSYPSFDDLEQLSENI 430

RESULT 9
S44096
enolpyruvylshikimate 3-phosphate synthase - Dichelobacter nodosus
C:Species: Dichelobacter nodosus
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 18-Jun-1999
A:Accession: S44096
R:Alm, R.A.; Dalrymple, B.P.; Mattick, J.S.
Submitted to the EMBL Data Library, January 1994
Description: Sequencing and expression of the ara gene from Dichelobacter nodosus.
Reference number: S44096
A:Accession: S44096
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-443 <ALM>
A:Cross-references: EMBL:Z29339; NID:g472931; PIDN:CAA82544.1; PID:g472932
C:Superfamily: 3-phosphoshikimate 1-carboxyvinyltransferase; 3-phosphoshikimate 1-carboxy
F:16-424/Domain: 3-phosphoshikimate 1-carboxyvinyltransferase homology <PSK>

Query Match 35.2%; Score 806; DB 2; Length 443;
Best Local Similarity 42.2%; Pred. No. 4.3e-44;
Matches 185; Conservative 74; Mismatches 169; Indels 10; Gaps 3;

QY 16 SGLSTVRIPGDKSISHRSFMEGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGD 75
Db 12 SALSGETITCGDKSNHALLAALAEQTEIRGFLACADCLATRAQLRALGVDIQREKE 71
QY 76 TWIDGVNGGLLAPEAPLDFGNATGRLTMGLVGVYDFDSTFTIGDASLTKRPMGRVLN 135
Db 72 IVTIRGVGLGLOPKAPLNMQNSGTSMLLAGILAAQRFESVLCGDESLEKRRPMQRIIT 131
QY 136 PLREMGVQVKSDEGDRPLVTLRGPKTPTTYRVPMAAQKSAVLLAGLNTPGIITVIE 195
Db 132 PLVQMGAKIVSHNSTAPLHISG-RPLTGIDYALPLPSAQLASCLILAGLLADGTYRLHT 190
QY 196 PLWTRDTEKMLQGFANLTVETDADGVRTIRLEGKGLTGQVIDVPGDPSSTAPFLVAA 255
191 CGISRDRHTERMLPFGGALEIKKE-----QIIIVTGGQKLHGCVLDIVGDLSSAAAPFMVAA 245
QY 256 LLVPGSDVTILNVLNPTRTGLILILQENGADIEVINPRLAGGEDVADLRVRSSTLKGVT 315
Db 246 LIAPRAEVIRNVGNPTRAATITLLQKMGRIELHQHGFWEAEPVADVIVYHSLKRGIT 305
QY 316 VPEDRAPSMIDRYPILAVAAAFEGATVNGLEELRVKESDRLSAVANGLKNGVDCDG 375
Db 306 VAPENIANAIDLPFFINAAACBEGTTFVGNLSSELRVKESDRLSAAQNLQILGVACDVG 365
QY 376 EFTSLVVRGRPDGKGLGNASGAATHLDHRIAMSLVGMGLVSENPTVDDATMIATSPFE 435
Db 366 ADFIHYGRSDRQFL---PARVNSFGDHRIAMSLAVAGVRAAGELLIDDGAVAAVMPQ 421
QY 436 FMDLMAGLCAKTELSDTK 453
Db 422 FRDFAAAGMNVGEKDAK 439

RESULT 10
D95159
3-phosphoshikimate 1-carboxyvinyltransferase [imported] - Streptococcus pneumoniae (stra
C:Species: Streptococcus pneumoniae
C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 24-Aug-2001
```

```
C:Accession: D95159
R:Tetzelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; H
on, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzappl
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morris
A:Title: Complete genome sequence of a virulent isolate of Streptococcus pneumoniae.
A:Reference number: A95000; MUID:21357209; PMID:11463916
A:Accession: D95159
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-427 <KUR>
A:Cross-references: GB:AE005672; PIDN:AAK75469.1; PID:g14972856; GSPDB:GN00164; TIGR:
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: SP1371
C:Superfamily: 3-phosphoshikimate 1-carboxyvinyltransferase; 3-phosphoshikimate 1-car
Query Match 35.1%; Score 803.5; DB 2; Length 427;
Best Local Similarity 42.8%; Pred. No. 5.9e-44;
Matches 184; Conservative 72; Mismatches 155; Indels 19; Gaps 7;

QY 18 LSGTVRIPGDKSISHRSFMEGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTW 77
Db 10 LHGSIKVPQDKSISHRSIIIFGSLAEGTKVYDILRGEDVLSTMQVFRDLGVEIEDKQVI 69
QY 78 IIDGVNGGLLAPEAPLDFGNATGRLTMGLVGVYDFDSTFTIGDASLTKRPMGRVLNPL 137
Db 70 TIQGVGMAGLKAPONALNMGNSGTSIRLSISGLAGADFEVEMFGDDSLSKRPMDRVTLPL 129
QY 138 REMGVQVKS-DCDRPLPVLIRGPKTPTTYRVPMAAQKSAVLLAGLNTPGIITVIEP 196
Db 130 KKMVSISQGTEDRLPRLKTKNLRPIHYELPIASQKSAFALQAKGESVIEEK 189
QY 197 IMTRDTEKMLQGFANLTVETDADGVRTIRLEGKGLTGQVIDVPGDPSSTAPFLVAAAL 256
Db 190 EYTRNTEKMLQGFHLSV----DG-KKIVGQPKLTKGVVFGDISSAAFWLAVGL 244
QY 257 LVPQSDVTILNVLNPTRTGLILILQENGADIEV--INPRLAGGEDVADLRVRSSTLKG 314
Db 245 IAPNSRLVLQNVGINETRTGIIDVIRAMGGKLEITEIDPVAKS---ATLIVESSDLKGT 300
QY 315 TVPEDRAPSMIDRYPILAVAAAFEGATVNGLEELRVKESDRLSAVANGLKNGVDCDE 374
Db 301 EIGGALIPRLIDELPIIALATOAQGVTVIKDAELKVKRETDRIQVVDALNSMGADITP 360
QY 375 GETSLVVRGRPDGKGLGNASGAATHLDHRIAMSLVGMGLVSENPTVDDATMIATSF 433
Db 361 TADGMIKCK-----SALHGARVNTFGDHRIGMNTAIAALLVADGEVELDRAEAINTSY 414
QY 434 PEFMDLMAGL 443
Db 415 PSFFDDLESIL 424

RESULT 11
S52580
3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) - Lactococcus lactis
N:Alternate names: 5-enolpyruvylshikimate-3-phosphate synthase
C:Species: Lactococcus lactis
C:Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 18-Jun-1999
A:Accession: S52580
R:Griffin, H.G.; Gasson, M.J.
Mol. Gen. Genet. 246, 119-127, 1995
A:Title: Genetic aspects of aromatic amino acid biosynthesis in Lactococcus lactis.
A:Reference number: S52579; MUID:95124293
A:Accession: S52580
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-430 <GR>
A:Cross-references: EMBL:X78413; NID:g683581; PIDN:CAA55180.1; PID:g683583
C:Superfamily: 3-phosphoshikimate 1-carboxyvinyltransferase; 3-phosphoshikimate 1-car
```


Thu Aug 15 13:49:58 2002

```

| | : |::||| |||| |::| : : | :||| ||||| : | : | |
Db 303 KPVKSPEEPTLIDEIPILAVLMAFADGVSEVKAGKELRYKESDRKAIVTNLRKLGVO 362
QY 372 CDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFVLMGLVSENPVTVDDATMIAT 431
Db 363 VEEFEDGFAIHGTKEIK-----GGVIETFKDHRIAMAFVGLVVEEVIIDHPECVTV 416
QY 432 SPPEFMD 438
Db 417 SYPEEWE 423
```

Search completed: August 15, 2002, 13:59:57
Job time: 131 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 15, 2002, 13:59:16 ; Search time 22.39 Seconds
(without alignments)
496.366 Million cell updates/sec

Title: US-09-464-099A-70

Perfect score: 2288

Sequence: 1 MLHGASSRPATARKSSGLSG.....FMDLMAGLAKIELSDTKAA 455

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents-AA:*
- 1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
 - 2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
 - 3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
 - 4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
 - 5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep:*
 - 6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2282	99.7	455	1	US-08-476-008-3
2	2282	99.7	455	1	US-08-306-063-3
3	2282	99.7	455	1	US-08-833-485-3
4	2282	99.7	455	4	US-09-137-440-3
5	2282	99.7	455	5	PCT-US91-06148A-3
6	1900.5	83.1	449	1	US-08-476-008-5
7	1900.5	83.1	449	1	US-08-476-008-7
8	1900.5	83.1	449	1	US-08-306-063-5
9	1900.5	83.1	449	1	US-08-306-063-7
10	1900.5	83.1	449	1	US-08-833-485-5
11	1900.5	83.1	449	1	US-08-833-485-7
12	1900.5	83.1	449	4	US-09-137-440-5
13	1900.5	83.1	449	4	US-09-137-440-7
14	1900.5	83.1	449	5	PCT-US91-06148A-5
15	1900.5	83.1	449	5	PCT-US91-06148A-7
16	867.5	37.9	447	1	US-08-476-008-67
17	867.5	37.9	447	1	US-08-306-063-67
18	867.5	37.9	447	1	US-08-833-485-67
19	867.5	37.9	447	4	US-09-137-440-67
20	806	35.2	443	1	US-08-476-008-69
21	806	35.2	443	1	US-08-306-063-69
22	806	35.2	443	1	US-08-833-485-69
23	806	35.2	443	4	US-09-137-440-69
24	803.5	35.1	427	1	US-09-325-881-2
25	798.5	34.9	427	2	US-08-896-345-2
26	798.5	34.9	427	4	US-09-226-091-2
27	795	34.3	428	1	US-08-476-008-42

28	785	34.3	428	1	US-08-306-063-42	Sequence 42, Appl
29	785	34.3	428	1	US-08-833-485-42	Sequence 42, Appl
30	785	34.3	428	4	US-09-137-440-42	Sequence 42, Appl
31	704.5	30.8	415	2	US-08-896-345-4	Sequence 4, Appl
32	704.5	30.8	415	4	US-09-226-091-4	Sequence 4, Appl
33	704.5	30.8	415	4	US-09-325-881-4	Sequence 4, Appl
34	577	25.2	430	1	US-08-476-008-44	Sequence 44, Appl
35	577	25.2	430	1	US-08-306-063-44	Sequence 44, Appl
36	577	25.2	430	1	US-08-833-485-44	Sequence 44, Appl
37	577	25.2	430	4	US-09-137-440-44	Sequence 44, Appl
38	402.5	17.6	427	1	US-08-476-008-60	Sequence 60, Appl
39	402.5	17.6	427	1	US-08-306-063-60	Sequence 60, Appl
40	402.5	17.6	427	1	US-08-833-485-60	Sequence 60, Appl
41	402.5	17.6	427	4	US-09-137-440-60	Sequence 60, Appl
42	378	16.5	427	4	US-09-243-374-7	Sequence 7, Appl
43	374.5	16.4	432	1	US-08-476-008-61	Sequence 61, Appl
44	374.5	16.4	432	1	US-08-306-063-61	Sequence 61, Appl
45	374.5	16.4	432	1	US-08-833-485-61	Sequence 61, Appl

ALIGNMENTS

RESULT 1
US-08-476-008-3
Sequence 3, Application US/08476008
Patent No. 5627061
GENERAL INFORMATION:
APPLICANT: Barry, Gerard F.
APPLICANT: Kishore, Ganesh M.
APPLICANT: Padgett, Stephen R.
APPLICANT: Stallings, William C.
TITLE OF INVENTION: Glyphosate Tolerant
5-Enolpyruvylshikimate-3-Phosphate Synthases
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BBAF
STREET: 700 Chesterfield Village Parkway
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63198
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,008
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/306,063
FILING DATE: 13-SEP-1994
APPLICATION NUMBER: US 07/749,611
FILING DATE: 28-AUG-1991
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/576,537
FILING DATE: 31-AUG-1990
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Hoerner Jr., Dennis R.
REGISTRATION NUMBER: 30,914
REFERENCE/DOCKET NUMBER: 38-21(106660)A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314)537-6099
TELEFAX: (314)537-6047
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 455 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
US-08-476-008-3

Query Match 99.7%: Score 2282; DB 1; Length 455;
Best Local Similarity 99.8%; Pred. No. 1.3e-202;
Matches 454; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSPFMFGGLASGETRITGLLEGEDVINTG 60
1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSPFMFGGLASGETRITGLLEGEDVINTG 60
61 KAMQAMGARIRKEGDTWIIDGVNGGLLAPEAPLDFGNAATGCRLTMLGVGYDFDSTFI 120
61 KAMQAMGARIRKEGDTWIIDGVNGGLLAPEAPLDFGNAATGCRLTMLGVGYDFDSTFI 120
121 GDASLTKRPMGRVNLPLNPLREMGVQVKSDEGDRPLVTLRGPKTPTITYRVPMAOQVSAV 180
121 GDASLTKRPMGRVNLPLNPLREMGVQVKSDEGDRPLVTLRGPKTPTITYRVPMAOQVSAV 180
181 LLAGLNTPGITTVIEPIIMTRDHTKMLQGFANLTVETDADGVRTIRLEGKLTGQVID 240
181 LLAGLNTPGITTVIEPIIMTRDHTKMLQGFANLTVETDADGVRTIRLEGKLTGQVID 240
241 VPGDPSSTAPFLVAALLVPGSDVTILNVLNMPRTGLIILTLQEMGADIEVINPRLAGGED 300
241 VPGDPSSTAPFLVAALLVPGSDVTILNVLNMPRTGLIILTLQEMGADIEVINPRLAGGED 300
301 VADLRVRSSTLKGTVVPEDRAPSMIDEYPIILAVAAFAEGATVMNGLEELRVKESDRLSA 360
301 VADLRVRSSTLKGTVVPEDRAPSMIDEYPIILAVAAFAEGATVMNGLEELRVKESDRLSA 360
361 VANGKLINGVDCDEGETSLVVRGPDGKGLGNASGAATAVATHLDHRIAMSFVLMGLVSENP 420
361 VANGKLINGVDCDEGETSLVVRGPDGKGLGNASGAATAVATHLDHRIAMSFVLMGLVSENP 420
421 VTVDATMTATSPFPMOLMAGLAKIELSDTKAA 455
421 VTVDATMTATSPFPMOLMAGLAKIELSDTKAA 455

RESULT 3
US-08-833-485-3
Sequence 3, Application US/08833485
Patent No. 5804425
GENERAL INFORMATION:
APPLICANT: Barry, Gerard F.
APPLICANT: Kishore, Ganesh M.
APPLICANT: Padgett, Stephen R.
APPLICANT: Stallings, William C.
TITLE OF INVENTION: Glyphosate Tolerant
TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F
STREET: 700 Chesterfield Village Parkway
CITY: St. Louis

MOLECULE TYPE: protein
US-08-476-008-3

Query Match 99.7%: Score 2282; DB 1; Length 455;
Best Local Similarity 99.8%; Pred. No. 1.3e-202;
Matches 454; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSPFMFGGLASGETRITGLLEGEDVINTG 60
1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSPFMFGGLASGETRITGLLEGEDVINTG 60
61 KAMQAMGARIRKEGDTWIIDGVNGGLLAPEAPLDFGNAATGCRLTMLGVGYDFDSTFI 120
61 KAMQAMGARIRKEGDTWIIDGVNGGLLAPEAPLDFGNAATGCRLTMLGVGYDFDSTFI 120
121 GDASLTKRPMGRVNLPLNPLREMGVQVKSDEGDRPLVTLRGPKTPTITYRVPMAOQVSAV 180
121 GDASLTKRPMGRVNLPLNPLREMGVQVKSDEGDRPLVTLRGPKTPTITYRVPMAOQVSAV 180
181 LLAGLNTPGITTVIEPIIMTRDHTKMLQGFANLTVETDADGVRTIRLEGKLTGQVID 240
181 LLAGLNTPGITTVIEPIIMTRDHTKMLQGFANLTVETDADGVRTIRLEGKLTGQVID 240
241 VPGDPSSTAPFLVAALLVPGSDVTILNVLNMPRTGLIILTLQEMGADIEVINPRLAGGED 300
241 VPGDPSSTAPFLVAALLVPGSDVTILNVLNMPRTGLIILTLQEMGADIEVINPRLAGGED 300
301 VADLRVRSSTLKGTVVPEDRAPSMIDEYPIILAVAAFAEGATVMNGLEELRVKESDRLSA 360
301 VADLRVRSSTLKGTVVPEDRAPSMIDEYPIILAVAAFAEGATVMNGLEELRVKESDRLSA 360
361 VANGKLINGVDCDEGETSLVVRGPDGKGLGNASGAATAVATHLDHRIAMSFVLMGLVSENP 420
361 VANGKLINGVDCDEGETSLVVRGPDGKGLGNASGAATAVATHLDHRIAMSFVLMGLVSENP 420
421 VTVDATMTATSPFPMOLMAGLAKIELSDTKAA 455
421 VTVDATMTATSPFPMOLMAGLAKIELSDTKAA 455

RESULT 2
US-08-306-063-3
Sequence 3, Application US/08306063
Patent No. 5633435
GENERAL INFORMATION:
APPLICANT: Barry, Gerard F.
APPLICANT: Kishore, Ganesh M.
APPLICANT: Padgett, Stephen R.
APPLICANT: Stallings, William C.
TITLE OF INVENTION: Glyphosate Tolerant
TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F
STREET: 700 Chesterfield Village Parkway
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63198
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/306,063
FILING DATE: 13-SEP-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,611
FILING DATE: 28-AUG-1991
CLASSIFICATION: 435

STATE: Missouri
 COUNTRY: USA
 ZIP: 63198
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/833,485
 FILING DATE: 07-APR-1997
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/306,063
 FILING DATE: 13-SEP-1994
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/749,611
 FILING DATE: 28-AUG-1991
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/576,537
 FILING DATE: 31-AUG-1990
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Hoerner Jr., Dennis R.
 REGISTRATION NUMBER: 30,914
 REFERENCE/DOCKET NUMBER: 38-21(15117)A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (314)737-6099
 TELEFAX: (314)737-6047
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 455 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-833-485-3

Query Match 99.7%; Score 2282; DB 1; Length 455;
 Best Local Similarity 99.8%; Pred. No. 1.3e-202;
 Matches 454; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLHGASSRATARKSSGLSGTVRIPGDKSISHSRSMFPGGLASGETRITGLLEGEDVINTG 60
 Db 1 MSHGASSRATARKSSGLSGTVRIPGDKSISHSRSMFPGGLASGETRITGLLEGEDVINTG 60
 61 KAMQAMGARIRKRGDGTWIIDGVNGGGLLAPEAPLDFGNATGCRLTWGLVGVYDFDSTFI 120
 Db 61 KAMQAMGARIRKRGDGTWIIDGVNGGGLLAPEAPLDFGNATGCRLTWGLVGVYDFDSTFI 120
 QY 121 GDASLTKRPMGRVNLPLREMGVQVSKEDGDRPLVTLRGPKTPPIYRVPMASAOQKSAV 180
 Db 121 GDASLTKRPMGRVNLPLREMGVQVSKEDGDRPLVTLRGPKTPPIYRVPMASAOQKSAV 180
 QY 181 LLAGLNTPGITTVIEPIMRDHTKMLQGFANLTVETDADGVRTIRLEGRKLTQVID 240
 Db 181 LLAGLNTPGITTVIEPIMRDHTKMLQGFANLTVETDADGVRTIRLEGRKLTQVID 240
 QY 241 VPGDPSSTAFFPLVAALLVPGSDVTILNVLNPNRTTGLITLQEMGADIEVINPRLAGGED 300
 Db 241 VPGDPSSTAFFPLVAALLVPGSDVTILNVLNPNRTTGLITLQEMGADIEVINPRLAGGED 300
 QY 301 VADLRVRSSTLKGVTVPDRAPSMIDEYPIILAVAAAFAGATVMNGLEELRVKESDRLSA 360
 Db 301 VADLRVRSSTLKGVTVPDRAPSMIDEYPIILAVAAAFAGATVMNGLEELRVKESDRLSA 360
 QY 361 VANGLKNGVDCDEGETSLVVRGRPGKGLGNASGAATAVATHDRIAMSLFVNGLVSENP 420
 Db 361 VANGLKNGVDCDEGETSLVVRGRPGKGLGNASGAATAVATHDRIAMSLFVNGLVSENP 420
 QY 421 VTVDDATMIATSPPEFMDLMAGLAKIELSDTKAA 455

Db 421 VTVDDATMIATSPPEFMDLMAGLAKIELSDTKAA 455
 RESULT 4
 US-09-137-440-3
 Sequence 3, Application US/09137440
 Patent No. 6248876
 GENERAL INFORMATION:
 APPLICANT: Barry, Gerard F.
 APPLICANT: Kishore, Ganesh M.
 APPLICANT: Padgett, Stephen R.
 APPLICANT: Stallings, William C.
 TITLE OF INVENTION: Glyphosate Tolerant
 NUMBER OF SEQUENCES: 69
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BBAF
 STREET: 700 Chesterfield Village Parkway
 CITY: St. Louis
 STATE: Missouri
 COUNTRY: USA
 ZIP: 63198
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/137,440
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/833,485
 FILING DATE: 07-APR-1997
 APPLICATION NUMBER: US 08/306,063
 FILING DATE: 13-SEP-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/749,611
 FILING DATE: 28-AUG-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/576,537
 FILING DATE: 31-AUG-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: Hoerner Jr., Dennis R.
 REGISTRATION NUMBER: 30,914
 REFERENCE/DOCKET NUMBER: 38-21(15117)A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (314)737-6099
 TELEFAX: (314)737-6047
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 455 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-137-440-3

Query Match 99.7%; Score 2282; DB 4; Length 455;
 Best Local Similarity 99.8%; Pred. No. 1.3e-202;
 Matches 454; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLHGASSRATARKSSGLSGTVRIPGDKSISHSRSMFPGGLASGETRITGLLEGEDVINTG 60
 Db 1 MSHGASSRATARKSSGLSGTVRIPGDKSISHSRSMFPGGLASGETRITGLLEGEDVINTG 60
 QY 61 KAMQAMGARIRKRGDGTWIIDGVNGGGLLAPEAPLDFGNATGCRLTWGLVGVYDFDSTFI 120
 Db 61 KAMQAMGARIRKRGDGTWIIDGVNGGGLLAPEAPLDFGNATGCRLTWGLVGVYDFDSTFI 120
 QY 121 GDASLTKRPMGRVNLPLREMGVQVSKEDGDRPLVTLRGPKTPPIYRVPMASAOQKSAV 180
 Db 121 GDASLTKRPMGRVNLPLREMGVQVSKEDGDRPLVTLRGPKTPPIYRVPMASAOQKSAV 180

Db 121 GDSLTRKPMGRVNLPLREMGVQKSEGDRLPVTLRGPKTPTITVYRVPMAQAQVKSAV 180
Qy 181 LLAGLNTPGITTVTIEPTMTDRHTEKMLQGGANLTVETDADGVRTIRLEGRGKLTGOVID 240
Db 181 LLAGLNTPGITTVTIEPTMTDRHTEKMLQGGANLTVETDADGVRTIRLEGRGKLTGOVID 240
Qy 241 VPGDPSSTAFPLVAALLVPGSDVTLNVLNMPTRTGLILTLQEMGADIEVINPRLAGGED 300
Db 241 VPGDPSSTAFPLVAALLVPGSDVTLNVLNMPTRTGLILTLQEMGADIEVINPRLAGGED 300
Qy 301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPIILAVAAAFAGATVMNGLEELRVKESDRLSA 360
Db 301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPIILAVAAAFAGATVMNGLEELRVKESDRLSA 360
Qy 361 VANGKLNGVDCDEGETSLVVRGPRDGKGLGNASGAATVATHLDRHRIAMSFVLMGLVSEN 420
Db 361 VANGKLNGVDCDEGETSLVVRGPRDGKGLGNASGAATVATHLDRHRIAMSFVLMGLVSEN 420
Qy 421 VTVDATMIATSFPEFMDLMAGLCAKIELSDTKAA 455
Db 421 VTVDATMIATSFPEFMDLMAGLCAKIELSDTKAA 455

RESULT 5
PCT-US91-06148A-3
; Sequence 3, Application PC/TUS9106148A
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Padgett, Stephen R.
; TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F
; STREET: 700 Chesterfield Village Parkway
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/06148A
; FILING DATE: 19910828
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/749,611
; FILING DATE: 28-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/576,537
; FILING DATE: 31-AUG-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Hoerner Jr., Dennis R.
; REGISTRATION NUMBER: 30,914
; REFERENCE/DOCKET NUMBER: 38-21(10535)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)537-6099
; TELEFAX: (314)537-6047
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 455 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US91-06148A-3

Query Match 99.7%; Score 2282; DB 5; Length 455;
Best Local Similarity 99.8%; Pred. No. 1.3e-202;

Matches 454; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSPFMFGGLASGETRITGLLEGEDVINTG 60
Db 1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSPFMFGGLASGETRITGLLEGEDVINTG 60
Qy 61 KAMQAMGARIRKEGDWTWIDGVNGGLLAPLDEPNAATGCRMTMGLVGVYDFDSTFI 120
Db 61 KAMQAMGARIRKEGDWTWIDGVNGGLLAPLDEPNAATGCRMTMGLVGVYDFDSTFI 120
Qy 121 GDASLTRKPMGRVNLPLREMGVQKSEGDRLPVTLRGPKTPTITVYRVPMAQAQVKSAV 180
Db 121 GDASLTRKPMGRVNLPLREMGVQKSEGDRLPVTLRGPKTPTITVYRVPMAQAQVKSAV 180
Qy 181 LLAGLNTPGITTVTIEPTMTDRHTEKMLQGGANLTVETDADGVRTIRLEGRGKLTGOVID 240
Db 181 LLAGLNTPGITTVTIEPTMTDRHTEKMLQGGANLTVETDADGVRTIRLEGRGKLTGOVID 240
Qy 241 VPGDPSSTAFPLVAALLVPGSDVTLNVLNMPTRTGLILTLQEMGADIEVINPRLAGGED 300
Db 241 VPGDPSSTAFPLVAALLVPGSDVTLNVLNMPTRTGLILTLQEMGADIEVINPRLAGGED 300
Qy 301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPIILAVAAAFAGATVMNGLEELRVKESDRLSA 360
Db 301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPIILAVAAAFAGATVMNGLEELRVKESDRLSA 360
Qy 361 VANGKLNGVDCDEGETSLVVRGPRDGKGLGNASGAATVATHLDRHRIAMSFVLMGLVSEN 420
Db 361 VANGKLNGVDCDEGETSLVVRGPRDGKGLGNASGAATVATHLDRHRIAMSFVLMGLVSEN 420
Qy 421 VTVDATMIATSFPEFMDLMAGLCAKIELSDTKAA 455
Db 421 VTVDATMIATSFPEFMDLMAGLCAKIELSDTKAA 455

RESULT 6
US-08-476-008-5
; Sequence 5, Application US/08476008
; Patent No. 5627061
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Padgett, Stephen R.
; APPLICANT: Stallings, William C.
; TITLE OF INVENTION: Glyphosate Tolerant
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F
; STREET: 700 Chesterfield Village Parkway
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,008
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/306,063
; FILING DATE: 13-SEP-1994
; APPLICATION NUMBER: US 07/749,611
; FILING DATE: 28-AUG-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/576,537
; FILING DATE: 31-AUG-1990
; CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Hoerner Jr., Dennis R.
REGISTRATION NUMBER: 30,914
REFERENCE/DOCKET NUMBER: 38-21(10660)A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314)537-6099
TELEFAX: (314)537-6047
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 449 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-476-008-5

Query Match 83.1%; Score 1900.5; DB 1; Length 449;
Best Local Similarity 82.9%; Pred. No. 2.2e-167;
Matches 373; Conservative 32; Mismatches 42; Indels 3; Gaps 1;
1 MLHGASSRATARKSSGLSGTVIRPGDKSISHSRPFGLASGETRITGLLEGEDVINTG 60
Db 1 MSHSASPKPATARRSEALTGEIRPGDKSISHSRPFGLASGETRITGLLEGEDVINTG 60
QY 61 KAMQAMGARIRKEGDWTIIDGVNGGGLLAPEAPLDFGNAATGRLTMGLVGVYDSDTFI 120
Db 61 RAMQAMGARIRKEGDWTIIDGVNGGGLLAPEALDFGNAATGRLTMGLVGVYDSDTFI 120
QY 121 GDSATKRPMPGRVLPNPLREMGVQVSEDDGRLPVTLRGPKTPTPIYRVPMSAQVKSAY 180
Db 121 GDSATKRPMPGRVLPNPLREMGVQVSEDDGRLPVTLRGPKTPTPIYRVPMSAQVKSAY 180
QY 181 LLAGLNTPGITTVIEPIVTRDHTKMLQFGANLTVETDADGVTRIRLEGRKLTGQVID 240
Db 181 LLAGLNTPGITTVIEPIVTRDHTKMLQFGANLTVETDADGVTRIRIRITGQGLVGTID 240
QY 241 VPGDSSSTAPFLVAALLVPGSDVTILNVLNMPTRTGLIILTLOEMGADIEVINPRLAGGED 300
Db 241 VPGDSSSTAPFLVAALLVPGSDVTIRNVLNMPTRTGLIILTLOEMGADIEVINARLAGGED 300
QY 301 VADLRVRSSTLKGTVPEDRAPSMIDEPYLAFAAFAEGATVMNGLEELRVKESDRLSA 360
Db 301 VADLRVRSSTLKGTVPEDRAPSMIDEPYLAFAAFAEGATVMNGLEELRVKESDRLSA 360
QY 361 VANGKLNGVDCDEGETSLVVRGPDGKGLGNASGAATVATHDHRITAMSLVGLVSENP 420
Db 361 VARGLEANGVDCDEGETSLVVRGPDGKGLGNASGAATVATHDHRITAMSLVGLVSENP 420
QY 421 VTVDATMIATSPFPEFMDLMAGLAKIELS 450
Db 418 VTVDSDNMIATSPFPEFMDMMPGLGAKIELS 447

RESULT 7
US-08-476-008-7
Sequence 7, Application US/08476008
Patent No. 5627061
GENERAL INFORMATION:
APPLICANT: Barry, Gerard F.
APPLICANT: Kishore, Ganesh M.
APPLICANT: Padgett, Stephen R.
APPLICANT: Stallings, William C.
TITLE OF INVENTION: Glycophate Tolerant
TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BBAF
STREET: 700 Chesterfield Village Parkway
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63198
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,008
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/306,063
FILING DATE: 13-SEP-1994
APPLICATION NUMBER: US 07/749,611
FILING DATE: 28-AUG-1991
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/576,537
FILING DATE: 31-AUG-1990
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Hoerner Jr., Dennis R.
REGISTRATION NUMBER: 30,914
REFERENCE/DOCKET NUMBER: 38-21(10660)A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314)537-6099
TELEFAX: (314)537-6047
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 449 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-476-008-7

Query Match 83.1%; Score 1900.5; DB 1; Length 449;
Best Local Similarity 82.9%; Pred. No. 2.2e-167;
Matches 373; Conservative 32; Mismatches 42; Indels 3; Gaps 1;
QY 1 MLHGASSRATARKSSGLSGTVIRPGDKSISHSRPFGLASGETRITGLLEGEDVINTG 60
Db 1 MSHSASPKPATARRSEALTGEIRPGDKSISHSRPFGLASGETRITGLLEGEDVINTG 60
QY 61 KAMQAMGARIRKEGDWTIIDGVNGGGLLAPEAPLDFGNAATGRLTMGLVGVYDSDTFI 120
Db 61 RAMQAMGARIRKEGDWTIIDGVNGGGLLAPEALDFGNAATGRLTMGLVGVYDSDTFI 120
QY 121 GDSATKRPMPGRVLPNPLREMGVQVSEDDGRLPVTLRGPKTPTPIYRVPMSAQVKSAY 180
Db 121 GDSATKRPMPGRVLPNPLREMGVQVSEDDGRLPVTLRGPKTPTPIYRVPMSAQVKSAY 180
QY 181 LLAGLNTPGITTVIEPIVTRDHTKMLQFGANLTVETDADGVTRIRLEGRKLTGQVID 240
Db 181 LLAGLNTPGITTVIEPIVTRDHTKMLQFGANLTVETDADGVTRIRIRITGQGLVGTID 240
QY 241 VPGDSSSTAPFLVAALLVPGSDVTILNVLNMPTRTGLIILTLOEMGADIEVINPRLAGGED 300
Db 241 VPGDSSSTAPFLVAALLVPGSDVTIRNVLNMPTRTGLIILTLOEMGADIEVINARLAGGED 300
QY 301 VADLRVRSSTLKGTVPEDRAPSMIDEPYLAFAAFAEGATVMNGLEELRVKESDRLSA 360
Db 301 VADLRVRSSTLKGTVPEDRAPSMIDEPYLAFAAFAEGATVMNGLEELRVKESDRLSA 360
QY 361 VANGKLNGVDCDEGETSLVVRGPDGKGLGNASGAATVATHDHRITAMSLVGLVSENP 420
Db 361 VARGLEANGVDCDEGETSLVVRGPDGKGLGNASGAATVATHDHRITAMSLVGLVSENP 420
QY 421 VTVDATMIATSPFPEFMDLMAGLAKIELS 450
Db 418 VTVDSDNMIATSPFPEFMDMMPGLGAKIELS 447

RESULT 8
US-08-306-063-5

Sequence 5, Application US/08306063
Patent No. 5633435
GENERAL INFORMATION:
APPLICANT: Barry, Gerard F.
APPLICANT: Kishore, Ganesh M.
APPLICANT: Padgett, Stephen R.
APPLICANT: Stallings, William C.
TITLE OF INVENTION: Glyphosate Tolerant
TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F
STREET: 700 Chesterfield Village Parkway
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63198
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/306.063
FILING DATE: 13-SEP-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,611
FILING DATE: 28-AUG-1991
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Hoerner Jr., Dennis R.
REGISTRATION NUMBER: 30,914
REFERENCE/DOCKET NUMBER: 38-21(10660)A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314)537-6099
TELEFAX: (314)537-6047
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 449 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-306-063-5

Query Match 83.1%; Score 1900.5; DB 1; Length 449;
Best Local Similarity 82.9%; Pred. No. 2.2e-167;
Matches 373; Conservative 32; Mismatches 42; Indels 3; Gaps 1;

QY 1 MLHGASRRPATARSSGLSTVRIPOGKSIHSRSMFGLASGETRITGLLEGEDVINTG 60
Db 1 MSHSASPKPATARRSEALTGEIRIPGKSIHSRSMFGLASGETRITGLLEGEDVINTG 60

QY 61 KAMQAMGAKIRKGGDWIDGVNGGLLAPLDFGNAATGCLTMGLVGVYDFDSTFI 120
Db 61 KAMQAMGAKIRKGGDWIINGVNGGLLQPEALDFGNAGTGARLTMLGLVGYDMKTSFI 120

QY 121 GDASLTMRPMGRVNLPLREMGVQKSEDDRLPVTILRGKTPPTIYRVPMSAQVKS 180
Db 121 GDASLTMRPMGRVNLPLREMGVQVEAADGRMPLTLIGPKTANPTIYRVPMSAQVKS 180

QY 181 LLAGLNTPGTITVIEPTMTDRHTEKMLQGGANLTVETDADGVRTIRLEGRKLTGOVID 240
Db 181 LLAGLNTPGTITVIEPTMTDRHTEKMLQGGADLTIVETDKDGVRRHRTITCGKLVGQITD 240

QY 241 VPGDPSSTAPPLVAALLVPGSDVTILNVLNPNRTTGLILTLQEMGADIEVINPLAGGED 300
Db 241 VPGDPSSTAPPLVAALLVPGSDVTIRNVLNPNRTTGLILTLQEMGADIEVLNRLAGGED 300

QY 301 VADLRVRSSTLKGVTVPEDRAPSMIDEPYILAVAAFAEGATVMNGLEELRVKESDRLSA 360
Db 301 VADLRVRSKLGKVVVPPERAPSMIDEPYVLAATAAFEGETVMDGLDELVRKESDRLAA 360

QY 361 VANGKLKINGVDCDEGETSLVVRGRPGKGLGNASGAATVATLHLDHRIAMSLFVNGLYSENP 420
Db 361 VARGLEANGVDCTEGEMSLTVRGRPGKGLG---GGTVATHLDHRIAMSLFVNGLYSENP 417

QY 421 VTVDATMTATSPPEFMDLMAGLCAKIELS 450
Db 418 VTVDDSNMIATSPPEFMDMMPGLCAKIELS 447

RESULT 9
US-08-306-063-7
Sequence 7, Application US/08306063
Patent No. 5633435
GENERAL INFORMATION:
APPLICANT: Barry, Gerard F.
APPLICANT: Kishore, Ganesh M.
APPLICANT: Padgett, Stephen R.
APPLICANT: Stallings, William C.
TITLE OF INVENTION: Glyphosate Tolerant
TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F
STREET: 700 Chesterfield Village Parkway
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63198
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/306.063
FILING DATE: 13-SEP-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,611
FILING DATE: 28-AUG-1991
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Hoerner Jr., Dennis R.
REGISTRATION NUMBER: 30,914
REFERENCE/DOCKET NUMBER: 38-21(10660)A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314)537-6099
TELEFAX: (314)537-6047
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 449 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-306-063-7

Query Match 83.1%; Score 1900.5; DB 1; Length 449;
Best Local Similarity 82.9%; Pred. No. 2.2e-167;
Matches 373; Conservative 32; Mismatches 42; Indels 3; Gaps 1;

QY 1 MLHGASRRPATARSSGLSTVRIPOGKSIHSRSMFGLASGETRITGLLEGEDVINTG 60
Db 1 MSHSASPKPATARRSEALTGEIRIPGKSIHSRSMFGLASGETRITGLLEGEDVINTG 60

```
; 61 KAMQAMGARIRKEDGTWIIDGVNGCLLAPAPLDGNAATGRLTMGLVGVYDFDSTFI 120
; :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 120
; 61 RAMQAMGAKIRKEDGVIIINGVNGCLLQPEAALDFGNAGTGARLTMLGLVGTYDMKTSFI 120
; :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 120
; 121 GDASTLRKPMGRVNLPLREMGVQKSEDDGRLPVTLRGPKTPTPIYRVPMSAQVKSVA 180
; :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 180
; 121 GDASTLRKPMGRVNLPLREMGVQEAADGDRMPLTLIGPKTANPIYRVPMSAQVKSVA 180
; :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 180
; 181 LLAGLNTPGTITVIEPIMTRDHTKMLQGFANLTVETDADGVRTIRLEGRKLTGQVID 240
; :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 240
; 181 LLAGLNTPGTITVIEPIMTRDHTKMLQGFADLTVETDKDGVHRHITGQGLVGTID 240
; :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 240
; 241 VPGDPSSTAPLVAALLVPGSDVTILNLMNPTRTGLITLQEMGADIEVINPLAGGED 300
; :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 300
; 241 VPGDPSSTAPLVAALLVPGSDVTIRNLMNPTRTGLITLQEMGADIEVINLARLAGGED 300
; :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 300
; 301 VADLRVRSSTLKGVTVPEDRAPSMIDEPYLAFAAFAEGATVMNGLEELRVKESDRLSA 360
; :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 360
; 301 VADLRVRSKLGKGVVPPERAPSMIDEPYLAFAAFAEGATVMNGLEELRVKESDRLSA 360
; :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 360
; 361 VANGKLNGVDCDEGETSLVVRGPRDCKGLGNASGAATVATHLDRHTAMSLVNGLVSENP 420
; :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 420
; 361 VARGLEANGVDCTEGEMSLTVRGRDCKGLG---GGTVATHLDRHTAMSLVNGLVSENP 420
; :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 420
; 421 VTVDATMIATSPPEFMDLMAGLAKIELS 450
; :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 450
; 418 VTVDNSMIATSPPEFMDMMPGLGAKIELS 447
; :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 447

RESULT 10
US-08-833-485-5
; Sequence 5, Application US/08833485
; Patent No. 5804425
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Padgett, Stephen R.
; APPLICANT: Stallings, William C.
; TITLE OF INVENTION: Glyphosate Tolerant
; TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/833,485
; FILING DATE: 07-APR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/306,063
; FILING DATE: 13-SEP-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,611
; FILING DATE: 28-AUG-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/576,537
; FILING DATE: 31-AUG-1990
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hoerner Jr., Dennis R.
; REGISTRATION NUMBER: 30,914
```

```
; REFERENCE/DOCKET NUMBER: 38-21(15117)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)737-6099
; TELEFAX: (314)737-6047
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 449 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-833-485-5

Query Match 83.1%; Score 1900.5; DB 1; Length 449;
Best Local Similarity 82.9%; Pred. No. 2.2e-167;
Matches 373; Conservative 32; Mismatches 42; Indels 3; Gaps 1;

QY 1 MLHGASSRPATARKSSGLSCTVRIPGDKSISHRSFMPGGLASGETRITGLLEGEDVINTG 60
; :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 60
; 1 MSHSASPKPATARRSEALTGEIRPGDKSISHRSFMPGGLASGETRITGLLEGEDVINTG 60
; :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 60
; 61 KAMQAMGARIRKEDGTWIIDGVNGCLLAPAPLDGNAATGRLTMGLVGVYDFDSTFI 120
; :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 120
; 61 RAMQAMGAKIRKEDGVIIINGVNGCLLQPEAALDFGNAGTGARLTMLGLVGTYDMKTSFI 120
; :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 120
; 121 GDASTLRKPMGRVNLPLREMGVQKSEDDGRLPVTLRGPKTPTPIYRVPMSAQVKSVA 180
; :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 180
; 121 GDASTLRKPMGRVNLPLREMGVQEAADGDRMPLTLIGPKTANPIYRVPMSAQVKSVA 180
; :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 180
; 181 LLAGLNTPGTITVIEPIMTRDHTKMLQGFANLTVETDADGVRTIRLEGRKLTGQVID 240
; :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 240
; 181 LLAGLNTPGTITVIEPIMTRDHTKMLQGFADLTVETDKDGVHRHITGQGLVGTID 240
; :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 240
; 241 VPGDPSSTAPLVAALLVPGSDVTILNLMNPTRTGLITLQEMGADIEVINPLAGGED 300
; :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 300
; 241 VPGDPSSTAPLVAALLVPGSDVTIRNLMNPTRTGLITLQEMGADIEVINLARLAGGED 300
; :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 300
; 301 VADLRVRSSTLKGVTVPEDRAPSMIDEPYLAFAAFAEGATVMNGLEELRVKESDRLSA 360
; :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 360
; 301 VADLRVRSKLGKGVVPPERAPSMIDEPYLAFAAFAEGATVMNGLEELRVKESDRLSA 360
; :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 360
; 361 VANGKLNGVDCDEGETSLVVRGPRDCKGLGNASGAATVATHLDRHTAMSLVNGLVSENP 420
; :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 420
; 361 VARGLEANGVDCTEGEMSLTVRGRDCKGLG---GGTVATHLDRHTAMSLVNGLVSENP 420
; :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 420
; 421 VTVDATMIATSPPEFMDLMAGLAKIELS 450
; :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 450
; 418 VTVDNSMIATSPPEFMDMMPGLGAKIELS 447
; :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 447
```

```
RESULT 11
US-08-833-485-7
; Sequence 7, Application US/08833485
; Patent No. 5804425
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Padgett, Stephen R.
; APPLICANT: Stallings, William C.
; TITLE OF INVENTION: Glyphosate Tolerant
; TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F
; STREET: 700 Chesterfield Village Parkway
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/833,485
FILING DATE: 07-APR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/306,063
FILING DATE: 13-SEP-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,611
FILING DATE: 28-AUG-1991
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/576,537
FILING DATE: 31-AUG-1990
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Hoerner Jr., Dennis R.
REGISTRATION NUMBER: 30,914
REFERENCE/DOCKET NUMBER: 38-21(15117)A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314)737-6099
TELEFAX: (314)737-6047
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 449 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-833-485-7

Query Match 83.1%; Score 1900.5; DB 1; Length 449;
Best Local Similarity 82.9%; Pred. No. 2.2e-167;
Matches 373; Conservative 32; Mismatches 42; Indels 3; Gaps 1;

QY 1 MLHGASSRPTARKSGSLGCTVRIPGDKSISHRSFPMFGGLASGETRITGLLEGEDVINTG 60
Db 1 MSHASPKPATARRSALGTEIRIPGDKSISHRSFPMFGGLASGETRITGLLEGEDVINTG 60
QY 61 KAMQAMGARIRKEGDWTIIDGVNGGGLLAPAPLDFGNATGCRLTMLGVGVDFDSTFI 120
Db 61 RAMQAMGAKIRKEDGVYIINGVNGCGLLOPEALDFGNAGTGARLTMLGVGVDFDSTFI 120
QY 121 GDASLTKRPMGRVNLPLREMGVQVKSEDDGRLPVTLRGPKTPTPIYRVPMSAQVKS AV 180
Db 121 GDASLSKRPNGRVNLPLREMGVQVVEAADGDRMPLTLIGPKTANPTIYRVPMSAQVKS AV 180
QY 181 LLAGLNTPGITTVIEPTIMTRDHTKMLQGFANLTVETDADGVRTIRLEGRKLTGQVID 240
Db 181 LLAGLNTPGITTVIEPTIMTRDHTKMLQGFADLTVDKDGVRHIRTGGKLVGQTID 240
QY 241 VPGDPSSTAPFLVAALLVPGSDVTILNVLNMPRTGLILTLQEMGADIEVINPRLAGGED 300
Db 241 VPGDPSSTAPFLVAALLVPGSDVTIRNVLNMPRTGLILTLQEMGADIEVINPRLAGGED 300
QY 301 VADLRVRSSTLKGVTPEDRAPSMIDEPYTLAVAAAFPAEGATVNGLEELRVKESORLSA 360
Db 301 VADLRVRASKLGVVPPERAPSMIDEPYVLAIAAFAEGATVMDGLDELVRVKSORLAA 360
QY 361 VANGKLKNGVDCDGETSLVVRPDKGKGLNAGSAAVATHLDRHIAVSLVWGLVSEN 420
Db 361 VARGLEANGVDCDGETSLVVRPDKGKGLG---GGTVATHLDRHIAVSLVWGLAAEK 417
QY 421 VTVDATMIATSPFEFMDLWAGLCAKIELS 450
Db 418 VTVDSDNNIATSPFEFMDMPGLGAKIELS 447

RESULT 12
US-09-137-440-5
; Sequence 5, Application US/09137440

Patent No. 6248876
GENERAL INFORMATION:
APPLICANT: Barry, Gerard F.
APPLICANT: Kishore, Ganesh M.
APPLICANT: Padgett, Stephen R.
APPLICANT: Stallings, William C.
TITLE OF INVENTION: Glycosylated Tolerant
TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F
STREET: 700 Chesterfield Village Parkway
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63198
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/137,440
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/833,485
FILING DATE: 07-APR-1997
APPLICATION NUMBER: US 08/306,063
FILING DATE: 13-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,611
FILING DATE: 28-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/576,537
FILING DATE: 31-AUG-1990
ATTORNEY/AGENT INFORMATION:
NAME: Hoerner Jr., Dennis R.
REGISTRATION NUMBER: 30,914
REFERENCE/DOCKET NUMBER: 38-21(15117)A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314)737-6099
TELEFAX: (314)737-6047
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 449 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-137-440-5

Query Match 83.1%; Score 1900.5; DB 4; Length 449;
Best Local Similarity 82.9%; Pred. No. 2.2e-167;
Matches 373; Conservative 32; Mismatches 42; Indels 3; Gaps 1;

QY 1 MLHGASSRPTARKSGSLGCTVRIPGDKSISHRSFPMFGGLASGETRITGLLEGEDVINTG 60
Db 1 MSHASPKPATARRSALGTEIRIPGDKSISHRSFPMFGGLASGETRITGLLEGEDVINTG 60
QY 61 KAMQAMGARIRKEGDWTIIDGVNGGGLLAPAPLDFGNATGCRLTMLGVGVDFDSTFI 120
Db 61 RAMQAMGAKIRKEDGVYIINGVNGCGLLOPEALDFGNAGTGARLTMLGVGVDFDSTFI 120
QY 121 GDASLTKRPMGRVNLPLREMGVQVKSEDDGRLPVTLRGPKTPTPIYRVPMSAQVKS AV 180
Db 121 GDASLSKRPNGRVNLPLREMGVQVVEAADGDRMPLTLIGPKTANPTIYRVPMSAQVKS AV 180
QY 181 LLAGLNTPGITTVIEPTIMTRDHTKMLQGFANLTVETDADGVRTIRLEGRKLTGQVID 240
Db 181 LLAGLNTPGITTVIEPTIMTRDHTKMLQGFADLTVDKDGVRHIRTGGKLVGQTID 240
QY 241 VPGDPSSTAPFLVAALLVPGSDVTILNVLNMPRTGLILTLQEMGADIEVINPRLAGGED 300

```

Db 241 VPGDPSAPPLVAALLVEGSDVTIRNVLNPNRTGILTLQEMGADIEVLNARLAGGED 300
QY 301 VADLRVRSSTLKGVTVPEDRAPSMIDEPYILAVAAFAEGATVYNGLEELRVKESDRLSA 360
Db 301 VADLRVRSSTLKGVTVPEDRAPSMIDEPYILAVAAFAEGATVYNGLEELRVKESDRLSA 360
QY 361 VANGKLINGVDCDEGETSLVVRGPDKGLGNAGSAAVATHLDRHRIAMSLVGLVSEN 420
Db 361 VARGLEANGVDCDEGETSLVVRGPDKGLGNAGSAAVATHLDRHRIAMSLVGLVSEN 420
QY 421 VTVDATMIATSPFMDLMAGLAKIELS 450
Db 418 VTVDSDNMIATSPFMDMMPGLGAKIELS 447

```

```

RESULT 13
US-09-137-440-7
; Sequence 7, Application US/09137440
; Patent No. 6248876
GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Padgett, Stephen R.
; APPLICANT: Stallings, William C.
; TITLE OF INVENTION: Glyphosate Tolerant
; TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/09/137,440
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/833,485
; FILING DATE: 07-APR-1997
; APPLICATION NUMBER: US 08/306,063
; FILING DATE: 13-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,611
; FILING DATE: 28-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/576,537
; FILING DATE: 31-AUG-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Hoerner Jr., Dennis R.
; REGISTRATION NUMBER: 30,914
; REFERENCE/DOCKET NUMBER: 38-21(15117)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)737-6099
; TELEFAX: (314)737-6047
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 449 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-137-440-7

```

```

Query Match 83.1%; Score 1900.5; DB 4; Length 449;
Best Local Similarity 82.9%; Pred. No. 2.2e-167;

```

```

Matches 373; Conservative 32; Mismatches 42; Indels 3; Gaps 1;
QY 1 MLHGASRPATARKSSGSLCTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
Db 1 MSHSASPATARRSEALTGEIRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
QY 61 KAMQAMGARIRKKGDTWIIDGVNGGGLLAPLAPLDFGNAAATGCRLTMLGLVGVYDFDSTFI 120
Db 61 RAMQAMGAKIRKKGDTWIIDGVNGGGLLAPLAPLDFGNAAATGCRLTMLGLVGVYDFDSTFI 120
QY 121 GDASLTKRPMGRVNLNPLREMGVQVKSSEDGRPLVTLRGPKTPTITVYRVPMAQAQVKSAY 180
Db 121 GDASLSKRPMPGRVNLNPLREMGVQVEAADGRMPLTLIGPKTANPITVYRVPMAQAQVKSAY 180
QY 181 LLAGLNTPGITTVIEPIWTRDHTKMLQGFGANLTVETDADGVTRIIEGRGKLTGQVID 240
Db 181 LLAGLNTPGITTVIEPIWTRDHTKMLQGFGANLTVETDADGVTRIIEGRGKLTGQVID 240
QY 241 VPGDPSAPPLVAALLVPGSDVTIRNVLNPNRTGILTLQEMGADIEVLNARLAGGED 300
Db 241 VPGDPSAPPLVAALLVPGSDVTIRNVLNPNRTGILTLQEMGADIEVLNARLAGGED 300
QY 301 VADLRVRSSTLKGVTVPEDRAPSMIDEPYILAVAAFAEGATVYNGLEELRVKESDRLSA 360
Db 301 VADLRVRSSTLKGVTVPEDRAPSMIDEPYILAVAAFAEGATVYNGLEELRVKESDRLSA 360
QY 361 VANGKLINGVDCDEGETSLVVRGPDKGLGNAGSAAVATHLDRHRIAMSLVGLVSEN 420
Db 361 VARGLEANGVDCDEGETSLVVRGPDKGLGNAGSAAVATHLDRHRIAMSLVGLVSEN 420
QY 421 VTVDATMIATSPFMDLMAGLAKIELS 450
Db 418 VTVDSDNMIATSPFMDMMPGLGAKIELS 447

```

```

RESULT 14
PCT-US91-06148A-5
; Sequence 5, Application PC/TUS9106148A
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Padgett, Stephen R.
; TITLE OF INVENTION: Glyphosate Tolerant
; TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: PCT/US91/06148A
; FILING DATE: 19910828
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/749,611
; FILING DATE: 28-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/576,537
; FILING DATE: 31-AUG-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Hoerner Jr., Dennis R.
; REGISTRATION NUMBER: 30,914
; REFERENCE/DOCKET NUMBER: 38-21(10535)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)537-6099

```

TELEFAX: (314)537-6047
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 449 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US91-06148A-5

Query Match 83.1%; Score 1900.5; DB 5; Length 449;
Best Local Similarity 82.9%; Pred. No. 2.2e-167;
Matches 373; Conservative 32; Mismatches 42; Indels 3; Gaps 1;

QY 1 MLHGASRPATARKSGSLGTVIRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
DB 1 MSHSASPKPATARRSEALTEIRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
QY 61 KAMQAMGARIRKEDGTWIIDGVNGGLLAPLDFGNATGCRLTMLGLVGVYDSDSTFI 120
DB 61 RAMQAMGARIRKEDGVIIINGVNGCLLOPEALDFGNAGTGARLTMLGLVGYTDMKTSFI 120
QY 121 GDASLTKRPMGRVNLPLREMGVQKSEDDGRLPVTILRGPKTPTITRYVPMSAQVKSAY 180
DB 121 GDASLTKRPMGRVNLPLREMGVQVEAADGDRMPLTLIGPKTANPITRYVPMSAQVKSAY 180
QY 181 LLAGLNTPGTITVIEPIMTRDHTKMLQGFAGNLTVETDADGVRTIRLEGRKLTGOVID 240
DB 181 LLAGLNTPGTITVIEPVMTRDHTKMLQGFADLTVDKDGVRHIRTGQGLVGGTID 240
QY 241 VPGDPSSTAPFLVAALLVPGSDVTILNVLNMPRTGLILTLQEMGADIEVINPRLAGGED 300
DB 241 VPGDPSSTAPFLVAALLVEGSDVTIRNVLNMPRTGLILTLQEMGADIEVLNARLAGGED 300
QY 301 VADLRVRSSTLKGVTVPEDRAPSMIDEPYILAVAAFAEGATVMNGLEELRVKESDRLSA 360
DB 301 VADLRVRSSTLKGVTVPEDRAPSMIDEPYILAVAAFAEGATVMNGLEELRVKESDRLSA 360
QY 361 VANGKLNGVDCDEGETSLVVRGRPDGKGLGNASGAATVATHDRIAMSLVNGLYSENP 420
DB 361 VARGLEANGVDCDEGETSLVVRGRPDGKGLG---GGTVATHDRIAMSLVNGLYSENP 420
QY 421 VTVDATMTATSPPEFMDLMAGLAKIELS 450
DB 418 VTVDSSNMATSPPEFMDMPGLGAKIELS 447

RESULT 15
T-US91-06148A-7
Sequence 7, Application PC/TUS9106148A
GENERAL INFORMATION:
APPLICANT: Barry, Gerard F.
APPLICANT: Kishore, Ganesh M.
APPLICANT: Padgett, Stephen R.
TITLE OF INVENTION: Glycophosphate Tolerant
TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F
STREET: 700 Chesterfield Village Parkway
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63198
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/06148A
FILING DATE: 19910828
CLASSIFICATION: 800

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/749,611
FILING DATE: 28-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/576,537
FILING DATE: 31-AUG-1990
ATTORNEY/AGENT INFORMATION:
NAME: Hoerner Jr., Dennis R.
REGISTRATION NUMBER: 30,914
REFERENCE/DOCKET NUMBER: 38-21(10535)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314)537-6099
TELEFAX: (314)537-6047
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 449 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US91-06148A-7

Query Match 83.1%; Score 1900.5; DB 5; Length 449;
Best Local Similarity 82.9%; Pred. No. 2.2e-167;
Matches 373; Conservative 32; Mismatches 42; Indels 3; Gaps 1;

QY 1 MLHGASRPATARKSGSLGTVIRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
DB 1 MSHSASPKPATARRSEALTEIRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
QY 61 KAMQAMGARIRKEDGTWIIDGVNGGLLAPLDFGNATGCRLTMLGLVGVYDSDSTFI 120
DB 61 RAMQAMGARIRKEDGVIIINGVNGCLLOPEALDFGNAGTGARLTMLGLVGYTDMKTSFI 120
QY 121 GDASLTKRPMGRVNLPLREMGVQKSEDDGRLPVTILRGPKTPTITRYVPMSAQVKSAY 180
DB 121 GDASLTKRPMGRVNLPLREMGVQVEAADGDRMPLTLIGPKTANPITRYVPMSAQVKSAY 180
QY 181 LLAGLNTPGTITVIEPIMTRDHTKMLQGFAGNLTVETDADGVRTIRLEGRKLTGOVID 240
DB 181 LLAGLNTPGTITVIEPVMTRDHTKMLQGFADLTVDKDGVRHIRTGQGLVGGTID 240
QY 241 VPGDPSSTAPFLVAALLVPGSDVTILNVLNMPRTGLILTLQEMGADIEVINPRLAGGED 300
DB 241 VPGDPSSTAPFLVAALLVEGSDVTIRNVLNMPRTGLILTLQEMGADIEVLNARLAGGED 300
QY 301 VADLRVRSSTLKGVTVPEDRAPSMIDEPYILAVAAFAEGATVMNGLEELRVKESDRLSA 360
DB 301 VADLRVRSSTLKGVTVPEDRAPSMIDEPYILAVAAFAEGATVMNGLEELRVKESDRLSA 360
QY 361 VANGKLNGVDCDEGETSLVVRGRPDGKGLGNASGAATVATHDRIAMSLVNGLYSENP 420
DB 361 VARGLEANGVDCDEGETSLVVRGRPDGKGLG---GGTVATHDRIAMSLVNGLYSENP 420
QY 421 VTVDATMTATSPPEFMDLMAGLAKIELS 450
DB 418 VTVDSSNMATSPPEFMDMPGLGAKIELS 447

Search completed: August 15, 2002, 13:59:17
Job time: 91 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 15, 2002, 13:57:46 ; Search time 22.39 Seconds

(without alignments)
496.366 Million cell updates/sec

Title: US-09-464-099A-3

Perfect score: 2288

Sequence: 1 MSHGASSRPATARKSSGLSG.....FMDLWAGLGAELSDTKAA 455

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 2442594 residues

Local number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2288	100.0	455	1	US-08-476-008-3
2	2288	100.0	455	1	US-08-306-063-3
3	2288	100.0	455	1	US-08-833-485-3
4	2288	100.0	455	4	US-09-137-440-3
5	2288	100.0	455	5	PCT-US91-06148A-3
6	1906.5	83.3	449	1	US-08-476-008-5
7	1906.5	83.3	449	1	US-08-476-008-7
8	1906.5	83.3	449	1	US-08-306-063-5
9	1906.5	83.3	449	1	US-08-306-063-7
10	1906.5	83.3	449	1	US-08-833-485-5
11	1906.5	83.3	449	1	US-08-833-485-7
12	1906.5	83.3	449	4	US-09-137-440-5
13	1906.5	83.3	449	4	US-09-137-440-7
14	1906.5	83.3	449	5	PCT-US91-06148A-5
15	1906.5	83.3	449	5	PCT-US91-06148A-7
16	867.5	37.9	447	1	US-08-476-008-67
17	867.5	37.9	447	1	US-08-306-063-67
18	867.5	37.9	447	1	US-08-833-485-67
19	867.5	37.9	447	4	US-09-137-440-67
20	806	35.2	443	1	US-08-476-008-69
21	806	35.2	443	1	US-08-306-063-69
22	806	35.2	443	1	US-08-833-485-69
23	806	35.2	443	4	US-09-137-440-69
24	803.5	35.1	427	4	US-09-325-881-2
25	798.5	34.9	427	2	US-08-896-345-2
26	798.5	34.9	427	4	US-09-226-091-2
27	785	34.3	428	1	US-08-476-008-42

28	785	34.3	428	1	US-08-306-063-42	Sequence 42, Appl
29	785	34.3	428	1	US-08-833-485-42	Sequence 42, Appl
30	785	34.3	428	4	US-09-137-440-42	Sequence 42, Appl
31	704.5	30.8	415	2	US-08-896-345-4	Sequence 4, Appl
32	704.5	30.8	415	4	US-09-226-091-4	Sequence 4, Appl
33	704.5	30.8	415	4	US-09-325-881-4	Sequence 4, Appl
34	577	25.2	430	1	US-08-476-008-44	Sequence 44, Appl
35	577	25.2	430	1	US-08-306-063-44	Sequence 44, Appl
36	577	25.2	430	1	US-08-833-485-44	Sequence 44, Appl
37	577	25.2	430	4	US-09-137-440-44	Sequence 44, Appl
38	402.5	17.6	427	1	US-08-476-008-60	Sequence 60, Appl
39	402.5	17.6	427	1	US-08-306-063-60	Sequence 60, Appl
40	402.5	17.6	427	1	US-08-833-485-60	Sequence 60, Appl
41	402.5	17.6	427	4	US-09-137-440-60	Sequence 60, Appl
42	378	16.5	427	4	US-09-243-374-7	Sequence 7, Appl
43	374.5	16.4	432	1	US-08-476-008-61	Sequence 61, Appl
44	374.5	16.4	432	1	US-08-306-063-61	Sequence 61, Appl
45	374.5	16.4	432	1	US-08-833-485-61	Sequence 61, Appl

ALIGNMENTS

RESULT 1
US-08-476-008-3
; Sequence 3, Application US/08476008
; Patent No. 5627061

; GENERAL INFORMATION:

; APPLICANT: Barry, Gerard F.

; APPLICANT: Kishore, Ganesh M.

; APPLICANT: Padgett, Stephen R.

; APPLICANT: Stallings, William C.

; TITLE OF INVENTION: Glyphosate Tolerant

; NUMBER OF SEQUENCES: 69

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F

; STREET: 700 Chesterfield Village Parkway

; CITY: St. Louis

; STATE: Missouri

; COUNTRY: USA

; ZIP: 63198

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/476,008

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/306,063

; FILING DATE: 13-SEP-1994

; APPLICATION NUMBER: US 07/749,611

; FILING DATE: 28-AUG-1991

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/576,537

; FILING DATE: 31-AUG-1990

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Hoerner Jr., Dennis R.

; REGISTRATION NUMBER: 30,914

; REFERENCE/DOCKET NUMBER: 38-21(10660)A

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (314)537-6099

; TELEFAX: (314)537-6047

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 455 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

MOLECULE TYPE: protein
US-08-476-008-3

Query Match 100.0%; Score 2288; DB 1; Length 455;
Best Local Similarity 100.0%; Pred. No. 2.6e-203; Indels 0; Gaps 0;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSHGASSRPTARKSGSLGTVRIIPGDKSISHSRPMFGGLASGETRITGLLEGEVDINTG 60
DB 1 MSHGASSRPTARKSGSLGTVRIIPGDKSISHSRPMFGGLASGETRITGLLEGEVDINTG 60

QY 61 KAMQAMGARIRKEGDTWIDGVNGGLLAPEAPLDFGNAATGCRLTMLGLVGVYDFDSTFI 120
DB 61 KAMQAMGARIRKEGDTWIDGVNGGLLAPEAPLDFGNAATGCRLTMLGLVGVYDFDSTFI 120

QY 121 GDASLTKRPMGRVNLPLREMGVQKSEDDGRLPVTLRGPKTPTPTITYRVPMSAQVKSAY 180
DB 121 GDASLTKRPMGRVNLPLREMGVQKSEDDGRLPVTLRGPKTPTPTITYRVPMSAQVKSAY 180

QY 181 LLAGLNTPGTITVIEPIIMTRDHTKMLQGFANLTVETDADGVTRTLRSGRGLTQVID 240
DB 181 LLAGLNTPGTITVIEPIIMTRDHTKMLQGFANLTVETDADGVTRTLRSGRGLTQVID 240

QY 241 VPGDPSSTAFPLVAALLVPGSDVTILNVLNPNTRTGLILTLQEMGADIEVINPRLAGGED 300
DB 241 VPGDPSSTAFPLVAALLVPGSDVTILNVLNPNTRTGLILTLQEMGADIEVINPRLAGGED 300

QY 301 VADLRVRSSTLKGTVTPEDRAPSMIDEYPLAVAAAFAGCATVMNGLEELRVKESDRLSA 360
DB 301 VADLRVRSSTLKGTVTPEDRAPSMIDEYPLAVAAAFAGCATVMNGLEELRVKESDRLSA 360

QY 361 VANGKLINGVDCDEGETSLVVRGPDGKGLGNASGAATHLDRHRIAMSLVNGLYSENP 420
DB 361 VANGKLINGVDCDEGETSLVVRGPDGKGLGNASGAATHLDRHRIAMSLVNGLYSENP 420

QY 421 VTVDATMIATSPFPEMDLMAGLAKIELSDTKAA 455
DB 421 VTVDATMIATSPFPEMDLMAGLAKIELSDTKAA 455

RESULT 2
US-08-063-3
Sequence 3, Application US/08306063
Patent No. 5633435
GENERAL INFORMATION:
APPLICANT: Barry, Gerard F.
APPLICANT: Kishore, Ganesh M.
APPLICANT: Padgett, Stephen R.
APPLICANT: Stallings, William C.
TITLE OF INVENTION: Glyphosate Tolerant
TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F
STREET: 700 Chesterfield Village Parkway
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63198
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/306,063
FILING DATE: 13-SEP-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,611
FILING DATE: 28-AUG-1991
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/576,537
FILING DATE: 31-AUG-1990
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Hoerner Jr., Dennis R.
REGISTRATION NUMBER: 30,914
REFERENCE/DOCKET NUMBER: 38-21(10660)A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314)537-6099
TELEFAX: (314)537-6047
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 455 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-306-063-3

Query Match 100.0%; Score 2288; DB 1; Length 455;
Best Local Similarity 100.0%; Pred. No. 2.6e-203; Indels 0; Gaps 0;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSHGASSRPTARKSGSLGTVRIIPGDKSISHSRPMFGGLASGETRITGLLEGEVDINTG 60
DB 1 MSHGASSRPTARKSGSLGTVRIIPGDKSISHSRPMFGGLASGETRITGLLEGEVDINTG 60

QY 61 KAMQAMGARIRKEGDTWIDGVNGGLLAPEAPLDFGNAATGCRLTMLGLVGVYDFDSTFI 120
DB 61 KAMQAMGARIRKEGDTWIDGVNGGLLAPEAPLDFGNAATGCRLTMLGLVGVYDFDSTFI 120

QY 121 GDASLTKRPMGRVNLPLREMGVQKSEDDGRLPVTLRGPKTPTPTITYRVPMSAQVKSAY 180
DB 121 GDASLTKRPMGRVNLPLREMGVQKSEDDGRLPVTLRGPKTPTPTITYRVPMSAQVKSAY 180

QY 181 LLAGLNTPGTITVIEPIIMTRDHTKMLQGFANLTVETDADGVTRTLRSGRGLTQVID 240
DB 181 LLAGLNTPGTITVIEPIIMTRDHTKMLQGFANLTVETDADGVTRTLRSGRGLTQVID 240

QY 241 VPGDPSSTAFPLVAALLVPGSDVTILNVLNPNTRTGLILTLQEMGADIEVINPRLAGGED 300
DB 241 VPGDPSSTAFPLVAALLVPGSDVTILNVLNPNTRTGLILTLQEMGADIEVINPRLAGGED 300

QY 301 VADLRVRSSTLKGTVTPEDRAPSMIDEYPLAVAAAFAGCATVMNGLEELRVKESDRLSA 360
DB 301 VADLRVRSSTLKGTVTPEDRAPSMIDEYPLAVAAAFAGCATVMNGLEELRVKESDRLSA 360

QY 361 VANGKLINGVDCDEGETSLVVRGPDGKGLGNASGAATHLDRHRIAMSLVNGLYSENP 420
DB 361 VANGKLINGVDCDEGETSLVVRGPDGKGLGNASGAATHLDRHRIAMSLVNGLYSENP 420

QY 421 VTVDATMIATSPFPEMDLMAGLAKIELSDTKAA 455
DB 421 VTVDATMIATSPFPEMDLMAGLAKIELSDTKAA 455

RESULT 3
US-08-833-485-3
Sequence 3, Application US/08833485
Patent No. 5804425
GENERAL INFORMATION:
APPLICANT: Barry, Gerard F.
APPLICANT: Kishore, Ganesh M.
APPLICANT: Padgett, Stephen R.
APPLICANT: Stallings, William C.
TITLE OF INVENTION: Glyphosate Tolerant
TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F
STREET: 700 Chesterfield Village Parkway
CITY: St. Louis

STATE: Missouri
COUNTRY: USA
ZIP: 63198
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/833,485
FILING DATE: 07-APR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/306,063
FILING DATE: 13-SEP-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,611
FILING DATE: 28-AUG-1991
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/576,537
FILING DATE: 31-AUG-1990
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Hoerner Jr., Dennis R.
REGISTRATION NUMBER: 30,914
REFERENCE/DOCKET NUMBER: 38-21(15117)A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314)737-6099
TELEFAX: (314)737-6047
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 455 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-833-485-3

Query Match 100.0%; Score 2288; DB 1; Length 455;
Best Local Similarity 100.0%; Pred. No. 2.6e-203;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSHGASSRATARKSSGLSGTVRIPGDKSISHRSMFGLASGETRITGLLEGEDVINTG 60
|||||
DB 1 MSHGASSRATARKSSGLSGTVRIPGDKSISHRSMFGLASGETRITGLLEGEDVINTG 60
|||||
61 KAMQAMGARIRKEGDTWIIDGVNGGLLAPEAPLDFGNAATGCRLTGMLGVGVYDFDSTFI 120
|||||
DB 61 KAMQAMGARIRKEGDTWIIDGVNGGLLAPEAPLDFGNAATGCRLTGMLGVGVYDFDSTFI 120
|||||
QY 121 GDASLTKRPMGRVLPNLRMGVQVKSSEDGDRPLVTILRGPKTPTITYRVPMSAQVKS AV 180
|||||
DB 121 GDASLTKRPMGRVLPNLRMGVQVKSSEDGDRPLVTILRGPKTPTITYRVPMSAQVKS AV 180
|||||
QY 181 LLAGLNTPGITTVIEPIWTRDHTERKMLQFGANLTVETDADGVRTIRLEGRKLTGQVID 240
|||||
DB 181 LLAGLNTPGITTVIEPIWTRDHTERKMLQFGANLTVETDADGVRTIRLEGRKLTGQVID 240
|||||
QY 241 VPGDPSSTAFFLVAALLVPGSDVTILNVLNPNTRGLIILTQEMGADIEVINPRLAGGED 300
|||||
DB 241 VPGDPSSTAFFLVAALLVPGSDVTILNVLNPNTRGLIILTQEMGADIEVINPRLAGGED 300
|||||
QY 301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPIAVAAFAEGATVMNGLEELRVKESDRLSA 360
|||||
DB 301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPIAVAAFAEGATVMNGLEELRVKESDRLSA 360
|||||
QY 361 VANGKLNGVDCDEGETSLVVRGRPDGKGLGNASGAATAHLDRHTAMSFVNGLVSENP 420
|||||
DB 361 VANGKLNGVDCDEGETSLVVRGRPDGKGLGNASGAATAHLDRHTAMSFVNGLVSENP 420
|||||
QY 421 VTVDATMTATSPFPMFMDLMAGLGAKIELSDTKAA 455
|||||

DB 421 VTVDATMTATSPFPMFMDLMAGLGAKIELSDTKAA 455
|||||
RESULT 4
US-09-137-440-3
Sequence 3, Application US/09137440
Patent No. 6248876
GENERAL INFORMATION:
APPLICANT: Barry, Gerard F.
APPLICANT: Kishore, Ganesh M.
APPLICANT: Padgett, Stephen R.
APPLICANT: Stallings, William C.
TITLE OF INVENTION: Glyphosate Tolerant
TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSER: Dennis R. Hoerner, Jr., Monsanto Co. BB4F
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63198
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/137,440
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/833,485
FILING DATE: 07-APR-1997
APPLICATION NUMBER: US 08/306,063
FILING DATE: 13-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,611
FILING DATE: 28-AUG-1991
APPLICATION NUMBER: US 07/576,537
FILING DATE: 31-AUG-1990
ATTORNEY/AGENT INFORMATION:
NAME: Hoerner Jr., Dennis R.
REGISTRATION NUMBER: 30,914
REFERENCE/DOCKET NUMBER: 38-21(15117)A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314)737-6099
TELEFAX: (314)737-6047
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 455 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-137-440-3

Query Match 100.0%; Score 2288; DB 4; Length 455;
Best Local Similarity 100.0%; Pred. No. 2.6e-203;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSHGASSRATARKSSGLSGTVRIPGDKSISHRSMFGLASGETRITGLLEGEDVINTG 60
|||||
DB 1 MSHGASSRATARKSSGLSGTVRIPGDKSISHRSMFGLASGETRITGLLEGEDVINTG 60
|||||
QY 61 KAMQAMGARIRKEGDTWIIDGVNGGLLAPEAPLDFGNAATGCRLTGMLGVGVYDFDSTFI 120
|||||
DB 61 KAMQAMGARIRKEGDTWIIDGVNGGLLAPEAPLDFGNAATGCRLTGMLGVGVYDFDSTFI 120
|||||
QY 121 GDASLTKRPMGRVLPNLRMGVQVKSSEDGDRPLVTILRGPKTPTITYRVPMSAQVKS AV 180
|||||

Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHSRPFMGGLASGETRITGLLEGEDVINTG 60
Db 1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHSRPFMGGLASGETRITGLLEGEDVINTG 60

Qy 61 KAMQAMGARIRKEGDTWIIDGVNGGGLLAPLDFGNAATGCRLLTMGLVGVYDFDSTFI 120
Db 61 KAMQAMGARIRKEGDTWIIDGVNGGGLLAPLDFGNAATGCRLLTMGLVGVYDFDSTFI 120

Qy 121 GDSLTTRKPMGRVNLNPLREMGVQVKSEGDRLPVTLRGPKTPTTITTYRVPNASAQVKSAV 180
Db 121 GDSLTTRKPMGRVNLNPLREMGVQVKSEGDRLPVTLRGPKTPTTITTYRVPNASAQVKSAV 180

Qy 181 LLAGLNTPGIITTVIEPIWTRDHTKMLQGFGANLTVEVDAGVTRIRLEGRKLTGQVID 240
Db 181 LLAGLNTPGIITTVIEPIWTRDHTKMLQGFGANLTVEVDAGVTRIRLEGRKLTGQVID 240

Qy 241 VPGDPSSTAFPLVAALLVPGSDVTILNVLNMPTRTGLTLTLOEMGADIEVINPRLAGGED 300
Db 241 VPGDPSSTAFPLVAALLVPGSDVTILNVLNMPTRTGLTLTLOEMGADIEVINPRLAGGED 300

Qy 301 VADLRVRSSTLKGTVVPEDRAPSMIDEPYILAVAAFAEGATVMNGLEELRVKESDRLSA 360
Db 301 VADLRVRSSTLKGTVVPEDRAPSMIDEPYILAVAAFAEGATVMNGLEELRVKESDRLSA 360

Qy 361 VANGKLNGVDCDEGETSLVVRGRPDGKGLGNASGAATAVATHLDHRIAMSFVLMGLVSNP 420
Db 361 VANGKLNGVDCDEGETSLVVRGRPDGKGLGNASGAATAVATHLDHRIAMSFVLMGLVSNP 420

Qy 421 VTVDATMTATSPFPEFMDLMAGLAKIELSDTKAA 455
Db 421 VTVDATMTATSPFPEFMDLMAGLAKIELSDTKAA 455

RESULT 6
US-08-476-008-5
; Sequence 5, Application US/08476008
; Patent No. 5627061
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Padgett, Stephen R.
; APPLICANT: Stallings, William C.
; TITLE OF INVENTION: Glyphosate Tolerant
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BBAF
; STREET: 700 Chesterfield Village Parkway
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,008
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/306,063
; FILING DATE: 13-SEP-1994
; APPLICATION NUMBER: US 07/749,611
; FILING DATE: 28-AUG-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/576,537
; FILING DATE: 31-AUG-1990
; CLASSIFICATION: 435

Db 121 GDSLTTRKPMGRVNLNPLREMGVQVKSEGDRLPVTLRGPKTPTTITTYRVPNASAQVKSAV 180
Qy 181 LLAGLNTPGIITTVIEPIWTRDHTKMLQGFGANLTVEVDAGVTRIRLEGRKLTGQVID 240
Db 181 LLAGLNTPGIITTVIEPIWTRDHTKMLQGFGANLTVEVDAGVTRIRLEGRKLTGQVID 240

Qy 241 VPGDPSSTAFPLVAALLVPGSDVTILNVLNMPTRTGLTLTLOEMGADIEVINPRLAGGED 300
Db 241 VPGDPSSTAFPLVAALLVPGSDVTILNVLNMPTRTGLTLTLOEMGADIEVINPRLAGGED 300

Qy 301 VADLRVRSSTLKGTVVPEDRAPSMIDEPYILAVAAFAEGATVMNGLEELRVKESDRLSA 360
Db 301 VADLRVRSSTLKGTVVPEDRAPSMIDEPYILAVAAFAEGATVMNGLEELRVKESDRLSA 360

Qy 361 VANGKLNGVDCDEGETSLVVRGRPDGKGLGNASGAATAVATHLDHRIAMSFVLMGLVSNP 420
Db 361 VANGKLNGVDCDEGETSLVVRGRPDGKGLGNASGAATAVATHLDHRIAMSFVLMGLVSNP 420

Qy 421 VTVDATMTATSPFPEFMDLMAGLAKIELSDTKAA 455
Db 421 VTVDATMTATSPFPEFMDLMAGLAKIELSDTKAA 455

RESULT 5
PCT-US91-06148A-3
; Sequence 3, Application PC/TUS9106148A
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Padgett, Stephen R.
; TITLE OF INVENTION: Glyphosate Tolerant
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BBAF
; STREET: 700 Chesterfield Village Parkway
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/06148A
; FILING DATE: 19910828
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/749,611
; FILING DATE: 28-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/576,537
; FILING DATE: 31-AUG-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Hoerner Jr., Dennis R.
; REGISTRATION NUMBER: 30,914
; REFERENCE/DOCKET NUMBER: 38-21(10535)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)537-6099
; TELEFAX: (314)537-6047
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 455 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US91-06148A-3

Query Match 100.0%; Score 2288; DB 5; Length 455;
Best Local Similarity 100.0%; Pred. No. 2.6e-203;

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,008
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/306,063
; FILING DATE: 13-SEP-1994
; APPLICATION NUMBER: US 07/749,611
; FILING DATE: 28-AUG-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/576,537
; FILING DATE: 31-AUG-1990
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hoerner Jr., Dennis R.
; REGISTRATION NUMBER: 30,914
; REFERENCE/DOCKET NUMBER: 38-21(10660)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)537-6099
; TELEFAX: (314)537-6047
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 449 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-476-008-7

```

```

APPLICATION NUMBER: US 01/516,531
FILING DATE: 31-AUG-1990
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Hoerner Jr., Dennis R.
REGISTRATION NUMBER: 30,914
REFERENCE/DOCKET NUMBER: 38-21(10660)A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314)537-6099
TELEFAX: (314)537-6047
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 449 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-476-008-7

Query Match      83.3%; Score 1906.5; DB 1; Length 449;
Best Local Similarity 83.1%; Pred. No. 4.7e-168;
Matches 374; Conservative 33; Mismatches 41; Indels 3; Gaps 1;

QY 1 MSHGASSRPATARKSSGSLGTVRIPGDKSIHRSFMFGGLASGETRITGLLEGEDVINTG 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 MSHSASPKPATARSEALTGEIRIPGDKSIHRSFMFGGLASGETRITGLLEGEDVINTG 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 61 KAMQAMGAKIRKEGDWTIIDGVGNGGLLAPAPLDFGNAATFCGLTMCGLVGCVYDFDSTFI 120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 RAMQAMGAKIRKEGDVWIIINGVNGCLLPAPALDFGNAGTIGARLTMCGLVCTYDMKTSFI 120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 121 GDASLTRPRMGVRVNLPRMGVQVKSGEDGRPLVTLRGPKPTPTITRVPMASAQVKSAV 180
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 GDASLSKRPMGVRVNLPRMGVQVYEAADGRMPLTLIGPKTANBITRVPMASAQVKSAV 180
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 181 LLASLNPFGTITVETPTWTRDHTKMLGGFGANLTVETDAGVTRIRLEGGRKLTGOVID 240
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

DB	181	LLAGLNTPGVTTVEPWNTRKDHTEKMLQGGGADULTVETURKDGVYKHRIITGQGKALVGQTID	241
Qy	241	VPGDPSSTAFPLVAALLVPGDSDVTILNVLNPNPTRTGLILTLOEMGADIIEVINPRLAGGED	300
DB	241	VPGDPSSTAFPLVAALLVEGSDVTIRNVLMNPRTGTGLILTLOEMGADIEVLNARLAGGED	300
Qy	301	VADLRVRSSLTGLVTPEDRAPSMIDEYPILAVAAAPAEGATVMNGLEBEURVKESDRLSA	360
DB	301	VADLRVRASKLKGVVPPERAPSMIDEYPYLATAASFAEGEYMDGLIDELRVKESDRLAA	360
Qy	361	VANCLKUNGDCDEGETSLVVRGRPDCKGLGNASGAATHLDHRTAMSPFLVNGLVSENP	420
DB	361	VARGLEANGVDCTEGEMSRLTVRGRPDCKGLUG---GGTVATHLDHRTAMSPFLVNGLAAREXP	417
Qy	421	VTVD DATMIATSFPFMDLMAGLCAKIELS	450
DB	418	VTVDDSNKIATSFPFMDMMPGLCAKIELS	447
RESULT 8 US-08-306-063-5			

; Sequence 5, Application US/08306063
; Patent No. 5633435
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Padgett, Stephen R.
; APPLICANT: Stallings, William C.
; TITLE OF INVENTION: Glyphosate Tolerant
; TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F
; STREET: 700 Chesterfield Village Parkway
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/306.063
; FILING DATE: 13-SEP-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,611
; FILING DATE: 28-AUG-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/576,537
; FILING DATE: 31-AUG-1990
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hoerner Jr., Dennis R.
; REGISTRATION NUMBER: 30,914
; REFERENCE/DOCKET NUMBER: 38-21(10660)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)537-6099
; TELEFAX: (314)537-6047
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 449 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-306-063-5

Query Match 83.3%; Score 1906.5; DB 1; Length 449;
Best Local Similarity 83.1%; Pred. No. 4.7e-168;
Matches 374; Conservative 32; Mismatches 41; Indels 3; Gaps 1;

QY 1 MSHGASRRPATAKSSGLSTVRIPGDKSIHRSFMFGGLASGETRITGLGEDVINTG 60
DB 1 MSHASPKPATARRSEALTGEIRIPGDKSIHRSFMFGGLASGETRITGLGEDVINTG 60
QY 61 KAMQAGCARLKEGDTWIDGVNGGLLAPLDFGNATGCRMTGLVGVYDFDSTFI 120
DB 61 RAMQAGAKIRKEGDTWIDGVNGGLLAPLDFGNATGCRMTGLVGVYDFDSTFI 120
QY 121 GDASLTKRPMGRVNLPLREMGVQKSDGDRPLVTILRGPKTPPTIYRVPMSAQVKS 180
DB 121 GDASLTKRPMGRVNLPLREMGVQKSDGDRPLVTILRGPKTPPTIYRVPMSAQVKS 180
QY 181 LLAGLNTPGTITVIEPTWTRDHEKMLQCGANLTVETDADGVRTIRLGRGKLTGQVID 240
DB 181 LLAGLNTPGTITVIEPTWTRDHEKMLQCGANLTVETDADGVRTIRLGRGKLTGQVID 240
QY 241 VPGDPSSTAFPLVAALLVPGSDVTILNLMNPRTGLILTLQEMGADIEVINPRLAGGED 300
DB 241 VPGDPSSTAFPLVAALLVPGSDVTILNLMNPRTGLILTLQEMGADIEVINPRLAGGED 300

QY 301 VADLRVRSSTIKGVTPEDRAPSMIDEPYILAVAAAFAGATVWNGLEELRVKESDRLSA 360
DB 301 VADLRVRSSTIKGVTPEDRAPSMIDEPYILAVAAAFAGATVWNGLEELRVKESDRLSA 360
QY 361 VANGKLINGVDCDEGETSLVVRGPRDCKGLGNASGAATAVATLHDHRIAMSLVWGLVSENP 420
DB 361 VARGLEANGVDCDEGETSLVVRGPRDCKGLGNASGAATAVATLHDHRIAMSLVWGLVSENP 420
QY 421 VTVDATMIATSPPEFMDLMAGLAKIELS 450
DB 421 VTVDATMIATSPPEFMDLMAGLAKIELS 450
QY 417 VTVDATMIATSPPEFMDLMAGLAKIELS 450
DB 417 VTVDATMIATSPPEFMDLMAGLAKIELS 450

RESULT 9
US-08-306-063-7
; Sequence 7, Application US/08306063
; Patent No. 5633435
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Padgett, Stephen R.
; APPLICANT: Stallings, William C.
; TITLE OF INVENTION: Glyphosate Tolerant
; TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F
; STREET: 700 Chesterfield Village Parkway
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/306.063
; FILING DATE: 13-SEP-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,611
; FILING DATE: 28-AUG-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/576,537
; FILING DATE: 31-AUG-1990
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hoerner Jr., Dennis R.
; REGISTRATION NUMBER: 30,914
; REFERENCE/DOCKET NUMBER: 38-21(10660)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)537-6099
; TELEFAX: (314)537-6047
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 449 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-306-063-7

Query Match 83.3%; Score 1906.5; DB 1; Length 449;
Best Local Similarity 83.1%; Pred. No. 4.7e-168;
Matches 374; Conservative 32; Mismatches 41; Indels 3; Gaps 1;

QY 1 MSHGASRRPATAKSSGLSTVRIPGDKSIHRSFMFGGLASGETRITGLGEDVINTG 60
DB 1 MSHASPKPATARRSEALTGEIRIPGDKSIHRSFMFGGLASGETRITGLGEDVINTG 60

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/833,485
FILING DATE: 07-APR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/306,063
FILING DATE: 13-SEP-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,611
FILING DATE: 28-AUG-1991
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/576,537
FILING DATE: 31-AUG-1990
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Hoerner Jr., Dennis R.
REGISTRATION NUMBER: 30,914
REFERENCE/DOCKET NUMBER: 38-21(15117)A
TELEPHONE: (314)737-6099
TELEFAX: (314)737-6047
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 449 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-833-485-7

Query Match 83.3%; Score 1906.5; DB 1; Length 449;
Best Local Similarity 83.1%; Pred. No. 4.7e-168;
Matches 374; Conservative 32; Mismatches 41; Indels 3; Gaps 1;
QY 1 MSHGASSRATARKSGLSCTVRIPGDKSISHRSFMEFGLASGETRITGLLEGEDVINTG 60
Db 1 MSHSASPKPATARRRSEALTEIRIPGDKSISHRSFMEFGLASGETRITGLLEGEDVINTG 60
QY 61 KAMQAMGARIRKEDGTWIDGVNGGLLAPPEALDFGNATGCRLTMLGVGVYDFDSTFI 120
Db 61 RAMQAMGAKIRKEDGVWIIINGVNGCLLQPEALDFGNATGCRLTMLGVGVYDMKTSFI 120
QY 121 GDASLTKRPMGRVNLPLREMGVQKSDGDRPLVTLRGPKTPPIYRVPMASQAQVKSAY 180
Db 121 GDASLSKRPNGRVNLPLREMGVQVEAADGDRMPLTLIGPKTANPIYRVPMASQAQVKSAY 180
QY 181 LLAGLNTPGTITVIEPTIMTRDHTKMLQGFANLTVETDADGVRTIRLEGRKLTQGVID 240
Db 181 LLAGLNTPGTITVIEPVMTRDHTKMLQGFADLTVDKDGVRHIRTGQGLVGGTID 240
QY 241 VPGDPSSTAPPLVAALLVPGSDVTILNLMNPTRTGLILTLOEMGADIEVINPRLAGGED 300
Db 241 VPGDPSSTAPPLVAALLVPGSDVTIRNLMNPTRTGLILTLOEMGADIEVINPRLAGGED 300
QY 301 VADLRVRSSTLKGTVTPEDRAPSMIDEPYTLAVAAFAEAGTVMNGLEELRVKESDRLSA 360
Db 301 VADLRVRSKLGTVTPEDRAPSMIDEPYVLAATAFAEAGTVMNGLEELRVKESDRLAA 360
QY 361 VANGKLUNGVCDEGETSVLVRRPGDKGLGNASGAARVATHLDRHRTAMSLVNGLVSENP 420
Db 361 VARGLEANGVDCTEGEMSLVRRPGDKGLG---GGTVATHLDRHRTAMSLVNGLVAAEKP 417
QY 421 VTVDATMIATSPPEFMDLMAGLGAKIELS 450
Db 418 VTVDSSNNIATSPPEFMDMPGLGAKIELS 447
RESULT 12
US-09-137-440-5
; Sequence 5, Application.US/09137440

Patent No. 6248876
GENERAL INFORMATION:
APPLICANT: Barry, Gerard F.
APPLICANT: Kishore, Ganesh M.
APPLICANT: Padgett, Stephen R.
APPLICANT: Stallings, William R.
TITLE OF INVENTION: Glyphosate Tolerant
5-Enolpyruvylshikimate-3-Phosphate Synthases
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F
STREET: 700 Chesterfield Village Parkway
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63198
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/137,440
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/833,485
FILING DATE: 07-APR-1997
APPLICATION NUMBER: US 08/306,063
FILING DATE: 13-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,611
FILING DATE: 28-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/576,537
FILING DATE: 31-AUG-1990
ATTORNEY/AGENT INFORMATION:
NAME: Hoerner Jr., Dennis R.
REGISTRATION NUMBER: 30,914
REFERENCE/DOCKET NUMBER: 38-21(15117)A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314)737-6099
TELEFAX: (314)737-6047
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 449 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-137-440-5

Query Match 83.3%; Score 1906.5; DB 4; Length 449;
Best Local Similarity 83.1%; Pred. No. 4.7e-168;
Matches 374; Conservative 32; Mismatches 41; Indels 3; Gaps 1;
QY 1 MSHGASSRATARKSGLSCTVRIPGDKSISHRSFMEFGLASGETRITGLLEGEDVINTG 60
Db 1 MSHSASPKPATARRRSEALTEIRIPGDKSISHRSFMEFGLASGETRITGLLEGEDVINTG 60
QY 61 KAMQAMGARIRKEDGTWIDGVNGGLLAPPEALDFGNATGCRLTMLGVGVYDFDSTFI 120
Db 61 RAMQAMGAKIRKEDGVWIIINGVNGCLLQPEALDFGNATGCRLTMLGVGVYDMKTSFI 120
QY 121 GDASLTKRPMGRVNLPLREMGVQKSDGDRPLVTLRGPKTPPIYRVPMASQAQVKSAY 180
Db 121 GDASLSKRPNGRVNLPLREMGVQVEAADGDRMPLTLIGPKTANPIYRVPMASQAQVKSAY 180
QY 181 LLAGLNTPGTITVIEPTIMTRDHTKMLQGFANLTVETDADGVRTIRLEGRKLTQGVID 240
Db 181 LLAGLNTPGTITVIEPVMTRDHTKMLQGFADLTVDKDGVRHIRTGQGLVGGTID 240
QY 241 VPGDPSSTAPPLVAALLVPGSDVTILNLMNPTRTGLILTLOEMGADIEVINPRLAGGED 300

Db	241	VPGDPSSTAPLVAALLVSGSDVIRNVLNPNPRTGLIITLOEMGADIEVLNARLGGD	300
Qy	301	VADLRVRSSTLKGVTVPEDRAPSNHIDEYPTLAVAAAFAEGATVYNGLEELRVKESDRLSA	360
Db	301	VADLRVRASLKLGVVPPERAPSNHIDEYPVLAATAAFAEGETVMDGLDELRVKESDRLSA	360
Qy	361	VANGLKLNGVDCDEGETSLVVRGPPGKGLGNASGAAVATHLDHRTAMSFVYGLVSEN	420
Db	361	VARGLEANGVDCTEGEMSLTVRGPDPKGLG---GGTVATHLDHRTAMSFVYGLAAEKP	417
Qy	421	VTVDATMIATSPFEFMDLMAGLGAKIELS	450
Db	418	VTVDDSNNIATSPFEFMDMMPLGAKIELS	447

RESULT 13
US-09-137-440-7
Sequence 7, Application US/09137440
Patent No. 6248876
GENERAL INFORMATION:
APPLICANT: Barry, Gerard F.
APPLICANT: Kishore, Ganesh M.
APPLICANT: Padgett, Stephen R.
APPLICANT: Stallings, William C.
TITLE OF INVENTION: Glyphosate Tolerant
TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. B44F
STREET: 700 Chesterfield Village Parkway
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63198

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/137,440
; FILING DATE:
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; PRIORITY APPLICATION NUMBER: US 08/833,485
; FILING DATE: 07-APR-1997
; APPLICATION NUMBER: US 08/306,063
; FILING DATE: 13-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,611
; FILING DATE: 28-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/576,537
; FILING DATE: 31-AUG-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Hoerner Jr., Dennis R.
; REGISTRATION NUMBER: 30,914
; REFERENCE/DOCKET NUMBER: 38-21(15117)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)737-6099
; TELEFAX: (314)737-6047
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 449 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-137-440-7

```

Query Match	83.38;	Score 1906.5;	DB 4;	Length 449;
Best Local Similarity	83.18;	Pred. No. 4.7e-168;		

[illegible]

RESULT 14

PCT/US91-061484-5

Sequence 5, Application PC/TUS91061484

GENERAL INFORMATION:

APPLICANT: Barry, Gerard F.

APPLICANT: Kishore, Ganesh M.

APPLICANT: Padgett, Stephen R.

TITLE OF INVENTION: Glyphosate Tolerant

TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases

NUMBER OF SEQUENCES: 36

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BBAF

STREET: 700 Chesterfield Village Parkway

CITY: St. Louis

STATE: Missouri

COUNTRY: USA

ZIP: 63198

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC Compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US91/061484

FILING DATE: 19910828

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/07/749,611

FILING DATE: 28-AUG-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/576,537

FILING DATE: 31-AUG-1990

ATTORNEY/AGENT INFORMATION:

NAME: Hoerner Jr., Dennis R.

REGISTRATION NUMBER: 30,914

REFERENCE/DOCKET NUMBER: 38-21(10535)

TELECOMMUNICATION INFORMATION:

TELEPHONE: (314)537-6099

TELEFAX: (314)537-6047
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 449 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US91-06148A-5

Query Match 83.3%; Score 1906.5; DB 5; Length 449;
Best Local Similarity 83.1%; Pred. No. 4.7e-168;
Matches 374; Conservative 32; Mismatches 41; Indels 3; Gaps 1;

QY 1 MSHGASSRPATARKSGSLGTVIRIPGDKSISHRSPFMFGGLASGETRITGLLEGEDVINTG 60
DB 1 MSHSASPATARRSEALTGETIRIPGDKSISHRSPFMFGGLASGETRITGLLEGEDVINTG 60
QY 61 KAMQAMGARIRKEGDTWIIDGVNGGGLLAPEALDFGNAATGCRLTMLGVGYDFDSTFI 120
DB 61 RAMQAMGAKIRKEGDTWIIDGVNGGGLLAPEALDFGNAATGCRLTMLGVGYDFDSTFI 120
QY 121 GDAISLTKRPMGRVNLPLREMGVQVKSDEGDRPLVTLRGPKTPTTYRVPVMAAQVKSAY 180
DB 121 GDAISLTKRPMGRVNLPLREMGVQVKSDEGDRPLVTLRGPKTPTTYRVPVMAAQVKSAY 180
QY 181 LLAGLNTPGTIVTIEPIMTRDHTKMLQGFAGNLVETDADGVRTIRLEGRKLTGQVID 240
DB 181 LLAGLNTPGTIVTIEPIMTRDHTKMLQGFAGNLVETDADGVRTIRLEGRKLTGQVID 240
QY 241 VPGDPSSTAFPLVAALLVPGSDVTILNVLNPNTRTGLILTQEMGADIEVINPRLAGGED 300
DB 241 VPGDPSSTAFPLVAALLVPGSDVTILNVLNPNTRTGLILTQEMGADIEVINPRLAGGED 300
QY 301 VADLRVRSSTLKGTVVPEDRAPSMIDEPILAVAAFAFAGATVMNGLEELRVKESDRLSA 360
DB 301 VADLRVRSSTLKGTVVPEDRAPSMIDEPILAVAAFAFAGATVMNGLEELRVKESDRLSA 360
QY 361 VANGKLINGVDCDEGETSLVVRGPDGKGLGNASGAATAVATHLDHRIAMSLVGLVSENP 420
DB 361 VANGKLINGVDCDEGETSLVVRGPDGKGLGNASGAATAVATHLDHRIAMSLVGLVSENP 420
QY 421 VTDDATMIATSPFPEMDMLMAGLAKIELS 450
DB 421 VTDDATMIATSPFPEMDMLMAGLAKIELS 450
QY 418 VTDDSNMIATSPFPEMDMMPGLGAKIELS 447
DB 418 VTDDSNMIATSPFPEMDMMPGLGAKIELS 447

Search completed: August 15, 2002, 13:59:16
Job time: 90 sec

TELEFAX: (314)537-6047
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 449 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US91-06148A-5

Query Match 83.3%; Score 1906.5; DB 5; Length 449;
Best Local Similarity 83.1%; Pred. No. 4.7e-168;
Matches 374; Conservative 32; Mismatches 41; Indels 3; Gaps 1;

QY 1 MSHGASSRPATARKSGSLGTVIRIPGDKSISHRSPFMFGGLASGETRITGLLEGEDVINTG 60
DB 1 MSHSASPATARRSEALTGETIRIPGDKSISHRSPFMFGGLASGETRITGLLEGEDVINTG 60
QY 61 KAMQAMGARIRKEGDTWIIDGVNGGGLLAPEALDFGNAATGCRLTMLGVGYDFDSTFI 120
DB 61 RAMQAMGAKIRKEGDTWIIDGVNGGGLLAPEALDFGNAATGCRLTMLGVGYDFDSTFI 120
QY 121 GDAISLTKRPMGRVNLPLREMGVQVKSDEGDRPLVTLRGPKTPTTYRVPVMAAQVKSAY 180
DB 121 GDAISLTKRPMGRVNLPLREMGVQVKSDEGDRPLVTLRGPKTPTTYRVPVMAAQVKSAY 180
QY 181 LLAGLNTPGTIVTIEPIMTRDHTKMLQGFAGNLVETDADGVRTIRLEGRKLTGQVID 240
DB 181 LLAGLNTPGTIVTIEPIMTRDHTKMLQGFAGNLVETDADGVRTIRLEGRKLTGQVID 240
QY 241 VPGDPSSTAFPLVAALLVPGSDVTILNVLNPNTRTGLILTQEMGADIEVINPRLAGGED 300
DB 241 VPGDPSSTAFPLVAALLVPGSDVTILNVLNPNTRTGLILTQEMGADIEVINPRLAGGED 300
QY 301 VADLRVRSSTLKGTVVPEDRAPSMIDEPILAVAAFAFAGATVMNGLEELRVKESDRLSA 360
DB 301 VADLRVRSSTLKGTVVPEDRAPSMIDEPILAVAAFAFAGATVMNGLEELRVKESDRLSA 360
QY 361 VANGKLINGVDCDEGETSLVVRGPDGKGLGNASGAATAVATHLDHRIAMSLVGLVSENP 420
DB 361 VANGKLINGVDCDEGETSLVVRGPDGKGLGNASGAATAVATHLDHRIAMSLVGLVSENP 420
QY 421 VTDDATMIATSPFPEMDMLMAGLAKIELS 450
DB 421 VTDDATMIATSPFPEMDMLMAGLAKIELS 450
QY 418 VTDDSNMIATSPFPEMDMMPGLGAKIELS 447
DB 418 VTDDSNMIATSPFPEMDMMPGLGAKIELS 447

RESULT 15
PCT-US91-06148A-7
Sequence 7, Application PC/TUS9106148A
GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Padgett, Stephen R.
; TITLE OF INVENTION: Glycophate Tolerant
; TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F
; STREET: 700 Chesterfield Village Parkway
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/06148A
; FILING DATE: 19910828
; CLASSIFICATION: 800

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
OM protein - protein search, using sw model
Run on: August 15, 2002, 14:05:43 ; Search time 51.67 Seconds
(without alignments)
1523.373 Million cell updates/sec

Title: US-09-464-099A-70
Perfect score: 2288
Sequence: 1 MLHGASSRPARATARKSSGLSG.....FMDLMAGLGAKIELSDTKAA 455

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 562222 seqs, 172994929 residues

11 number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SPREMBL19:
2: sp_archaea:
3: sp_bacteria:
4: sp_fungi:
5: sp_human:
6: sp_invertebrate:
7: sp_mammal:
8: sp_mhc:
9: sp_organelle:
10: sp_plant:
11: sp_rodent:
12: sp_virus:
13: sp_vertebrate:
14: sp_unclassified:
15: sp_rvirus:
16: sp_bacteriap:
17: sp_archaeap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES						Description	
Result No.	Score	Query Match	Length	DB ID			
1	2098	91.7	455	16	Q92SV5	Q92SV5 rhizobium m	
2	1855.5	81.1	480	2	Q9AGV2	Q9AGV2 brucella ab	
3	1783	77.9	452	16	Q98CC1	Q98CC1 rhizobium l	
4	1135	49.6	443	16	Q9A2H2	Q9A2H2 caulobacter	
5	989.5	43.2	746	16	Q9H269	Q9H269 pseudomonas	
6	912.5	39.9	454	16	Q9PB21	Q9PB21 xylella fas	
7	884	38.6	431	16	Q9KCA6	Q9KCA6 bacillus ha	
8	820	35.8	430	16	Q99Z83	Q99Z83 streptococ	
9	815.5	35.6	428	16	Q92A85	Q92A85 listeria in	
10	781.5	34.2	428	2	Q9ANY6	Q9ANY6 enterococu	
11	610	26.7	432	16	Q99U25	Q99U25 staphylococ	
12	470.5	20.6	207	2	Q9RH28	Q9RH28 pseudomonas	
13	470.5	20.6	439	17	Q9HQCI	Q9HQCI halobacteri	
14	416.5	18.2	440	2	Q9L213	Q9L213 streptomyce	
15	401.5	17.5	426	16	Q9KR80	Q9KR80 vibrio chol	
16	379	16.6	427	2	Q93ED4	Q93ED4 yersinia ru	

17	351.5	15.4	428	16	Q97KM2	Q97km2 clostridium
18	332.5	14.5	438	2	Q9K4A7	Q9K4a7 streptomyce
19	324.5	14.2	447	16	Q9K9D5	Q9K9d5 bacillus ha
20	306	13.4	516	10	Q946V0	Q946v0 dicliptera
21	302	13.2	521	10	Q9FVP6	Q9Fvp6 arabidopsis
22	299	13.1	444	10	Q24566	Q24566 zea mays (m
23	298	13.0	516	10	Q946U9	Q946u9 dicliptera
24	297.5	13.0	511	10	Q93VK6	Q93vk6 oryza sativ
25	293.5	12.8	445	8	Q95AK0	Q95ak0 eleusine in
26	292.5	12.8	445	8	Q95AK1	Q95ak1 eleusine in
27	285	12.5	433	16	Q9JTT3	Q9Jtt3 neisseria m
28	280	12.2	414	17	Q980I5	Q980i5 sulfolobus
29	275	12.0	433	16	Q9JYU1	Q9Jyul neisseria m
30	261	11.4	408	17	Q96Y91	Q96y91 sulfolobus
31	260.5	11.4	427	17	Q9YC47	Q99eq0 aeropyrum p
32	229	10.0	332	2	Q9ZEQ0	Q99eq0 actinobacil
33	223	9.7	410	17	Q978S3	Q978s3 thermoplasm
34	216.5	9.5	391	10	Q80428	Q80428 oryza sativ
35	206.5	9.0	448	2	Q91LU5	Q91lu5 streptomyce
36	203.5	8.9	446	2	Q9S0N1	Q99278 streptococ
37	185.5	8.1	419	16	Q99278	Q99278 streptococ
38	184.5	8.1	417	16	Q9K1Q9	Q9K1q9 neisseria m
39	184	8.0	418	16	Q97DD9	Q97dd9 clostridium
40	183.5	8.0	417	16	Q9JWS7	Q9Jws7 neisseria m
41	182	8.0	428	16	Q9K6E5	Q9K6e5 bacillus ha
42	181	7.9	419	2	Q9S5W5	Q9S5w5 escherichia
43	179	7.8	347	10	Q9AT37	Q9at37 lolium rigi
44	179	7.8	423	16	Q927U1	Q927u1 listeria in
45	178.5	7.8	417	2	Q9EXE3	Q9exe3 mycobacteri

ALIGNMENTS

RESULT 1

Q92SV5 PRELIMINARY; PRT; 455 AA.
AC Q92SV5;
DT 01-DEC-2001 (TREMREL. 19, Created)
DT 01-DEC-2001 (TREMREL. 19, Last sequence update)
DT 01-DEC-2001 (TREMREL. 19, Last annotation update)
DE PUTATIVE 3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTANSFERASE PROTEIN (EC 2.5.1.19).
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21368234; PubMed=11474104;
RA Galibert F., Finan T.M., Long S.R., Puehler A., Abola P., Ampe F., Barloy-Hubler F., Barnett M.J., Becker A., Boistard P., Bothe G., Boutry M., Bowser L., Buhrmester J., Cadieu E., Capela D., Chain P., Cowie A., Davis R.W., Dreano S., Federspiel N.A., Fisher R.F., Gloux S., Godrie T., Goffeau A., Golding B., Gouzy J., Gurjal M., Hernandez-Lucas I., Hong A., Huizar L., Hyman R.W., Jones T., Kahn D., Kahn M.L., Kalman S., Keating D.H., Kiss E., Komp C., Lelaure V., Masuy D., Palm C., Peck M.C., Pohl T.M., Portetelle D., Purnelle B., Ramsperger U., Surzycki R., Thebault P., Vandenbol M., Vorhoelter F.J., Weidner S., Wells D.H., Wong K., Yeh K.-C., Batut J.;
RT "The composite genome of the legume symbiont Sinorhizobium meliloti.";
RL Science 293:668-672(2001).
DR EMBL; AL591783; CAC41690.1; -;
KW Transferrase; Complete proteome.
SQ SEQUENCE 455 AA; 47900 MW; 97659E1C7E1021B5 CRC64;

Query Match 91.7%; Score 2098; DB 16; Length 455;
Best Local Similarity 90.5%; Pred. No. 2.3e-119;
Matches 412; Conservative 18; Mismatches 25; Indels 0; Gaps 0;

Qy 1 MLHGASSRPARATARKSSGLSGTVRIPGDKSISHRSFPMFGGLASGETRITGLLGEDVINTG 60

```

|:::||||| |:::||||| |:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Db 1 MSHGNRPATARKSSDDKGLTIRIPGDKSISHRSFMFGGLAAGETRITGLLEGEDVINTG 60
QY 61 KAMQAMGARIRKEGDTWIIDGVNGGLLAPAPLDFGNAAATGCRLTMLGVGVYDFDSTFI 120
Db 61 KAMQAMGARIRKEGDTWIIDGVNGGLLAPAPLDFGNAGTGCRLTMGLGVYDFDSTFI 120
QY 121 GDASLTKRPMGRVNLPLREMGVQVKSDEGDRPLPVTLRGPKPTPTIYRVPMSAQVKS AV 180
Db 121 GDASLTKRPMGRVNLPLREMGVQVKSDEGDRPLPVTLRGPKPTPTIYRVPMSAQVKS AV 180
QY 181 LLAGLNTPGITTVIEPIIMTRDHTKMLQGFANLTVEVDAGVTRIRLEGRGKLTGQVID 240
Db 181 LLAGLNTPGITTVIEPVVTRDHTKMLQGFANLTVEVDAGVTRIRLEGRGKLTGQVID 240
QY 241 VPGDPSSTAFPLVAALLVPGSDVITLNLVLMNPTRTGLTLTLOEMGADIEVINPRLAGGED 300
Db 241 VPGDPSSTAFPLVAGLIVPGSDITLNLVLMNPTRTGLTLTLOEMGANTEVMNKRLLAGGED 300
QY 301 VADLRVRSSTLKGTVVPEDRAPSMIDEPYILAVAAFAEGATVMNGLEELRVKESDRLSA 360
Db 301 VADLRVRSSTLKGTVVPEDRAPSMIDEPYILAVAAFAEGATVMNGLEELRVKESDRLSA 360
QY 361 VANGKLKNGVDCDEGETSLVVRGRPDGKGLGNASGAATVATHLDHRIAMSFVLMGLVSEN 420
Db 361 VADGLKNGVDCDEGEASLVVRGRPGGKGLGKISGGQVKTLDHRIAMSFVLMGLASEHP 420
QY 421 VTVDATMIATSPFPEFMDLMAGLAKIELSDTKAA 455
Db 421 VTVDATMIATSPFPEFMDLMAGLAKIEAENKAA 455
```

```

RESULT 2
Q9AGV2 PRELIMINARY; PRT; 480 AA.
AC Q9AGV2;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE 5-ENOLPYRUVYL SHIKIMATE 3-PHOSPHATE SYNTHASE (EC 2.5.1.19).
GN AROA.
OS Brucella abortus.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=235;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2308;
RL "Characterization of the aroA gene of Brucella abortus and construction of an aroA mutant.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF326475; AAK27445.1; -.
DR InterPro; IPR001986; EPSP_synthase.
DR Pfam; PF00275; EPSP_synthase; 1.
DR ProDom; PD001867; EPSP_synthase; 1.
DR PROSITE; PS00104; EPSP_SYNTHASE_1; UNKNOWN_1.
DR PROSITE; PS00885; EPSP_SYNTHASE_2; UNKNOWN_1.
KW Transferase.
SQ SEQUENCE 480 AA; 50634 MW; 3D55323944DA8C91 CRC64;
```

```

Query Match 81.1%; Score 1855.5; DB 2; Length 480;
Best Local Similarity 81.4%; Pred. No. 1.2e-104;
Matches 364; Conservative 35; Mismatches 45; Indels 3; Gaps 1;

QY 1 MLHCASRRPATARKSSGLTGVIPGDKSISHRSFMFGGLASCTETRTITGLLEGEDVINTG 60
Db 31 MSHSACPAPATARKSQALETGIRIPGDKSISHRSFMFGGLASCTETRTITGLLEGEDVINTG 90
QY 61 KAMQAMGARIRKEGDTWIIDGVNGGLLAPAPLDFGNAAATGCRLTMLGVGVYDFDSTFI 120
Db 91 KAMQAMGARIRKEGDTWIIDGVNGGLLAPAPLDFGNAGTGCRLTMGLGVYDFDSTFI 150
```

```

QY 121 GDASLTKRPMGRVNLPLREMGVQVKSDEGDRPLPVTLRGPKPTPTIYRVPMSAQVKS AV 180
Db 151 GDASLTKRPMGRVNLPLREMGVQVKSDEGDRPLPVTLRGPKPTPTIYRVPMSAQVKS AV 210
QY 181 LLAGLNTPGITTVIEPIIMTRDHTKMLQGFANLTVEVDAGVTRIRLEGRGKLTGQVID 240
Db 211 LLAGLNTPGITTVIEPVVTRDHTKMLQGFADLTVEVDAGVTRIRIVGOGKLTGQVID 270
QY 241 VPGDPSSTAFPLVAALLVPGSDVITLNLVLMNPTRTGLTLTLOEMGADIEVINPRLAGGED 300
Db 271 VPGDPSSTAFPLVALLVPGSEVTRINVLNMPRTGLTLTLOEMGADIEIIDPRLAGGED 330
QY 301 VADLRVRSSTLKGTVVPEDRAPSMIDEPYILAVAAFAEGATVMNGLEELRVKESDRLSA 360
Db 331 VADLRVRSSTLKGTVVPEDRAPSMIDEPYILAVAAFAEGATVMNGLEELRVKESDRLSA 390
QY 361 VANGKLKNGVDCDEGETSLVVRGRPDGKGLGNASGAATVATHLDHRIAMSFVLMGLVSEN 420
Db 391 VARGLEANGVDCDEGETSLVVRGRPGGKGLG--GGTVATHLDHRIAMSFVLMGLASEKP 447
QY 421 VTVDATMIATSPFPEFMDLMAGLAKI 447
Db 448 VTVDATMIATSPFPEFMDLMAGLAKI 474

RESULT 3
Q98CC1 PRELIMINARY; PRT; 452 AA.
AC Q98CC1;
DT 01-OCT-2001 (TREMBLrel. 18, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE 3-PHOSPHOSHIMIKIMATE 1-CARBOXYVINYLTRANSFERASE.
GN MLL5213.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RC MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Iidesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL; AP003006; BAB51700.1; -.
DR InterPro; IPR001986; EPSP_synthase.
DR Pfam; PF00275; EPSP_synthase; 1.
DR ProDom; PD001867; EPSP_synthase; 1.
DR Transferase; Complete proteome.
KW Transferase.
SQ SEQUENCE 452 AA; 47455 MW; 2B52983E3523B938 CRC64;
```

```

Query Match 77.9%; Score 1783; DB 16; Length 452;
Best Local Similarity 77.8%; Pred. No. 2.7e-100;
Matches 351; Conservative 33; Mismatches 63; Indels 4; Gaps 2;
```

```

QY 1 MLHCASRRPATARKSSGLTGVIPGDKSISHRSFMFGGLASCTETRTITGLLEGEDVINTG 60
Db 1 MSHAAAKPATARKSQALETGIRIPGDKSISHRSFMFGGLASCTETRTITGLLEGEDVINTG 60
QY 61 KAMQAMGARIRKEGDTWIIDGVNGGLLAPAPLDFGNAAATGCRLTMLGVGVYDFDSTFI 120
Db 61 ANKAMGAHIEKRAEAWIRGTGNGALLQEPGLDFGNAGTGSRLTMGLGVYDMETTFI 120
QY 121 GDASLTKRPMGRVNLPLREMGVQV--KSEGDRLPVTLRGPKPTPTIYRVPMSAQVKS AV 179
Db 121 GDASLTKRPMGRVNLPLREMGVQV--KSEGDRLPVTLRGPKPTPTIYRVPMSAQVKS AV 179
```

Db 121 GDASLSGRPMGRVLEPLRQMGVQLKATPGDRMPITLHGPKHAAPIYRVPMASAOVKS 180

QY 180 VLLAGLNTPGITTVIEPIIMTRDHTKMLQGFQGANITVETDADGVRTIRLEGRGKLTQGV 239

Db 181 VLLAGLNTPGITTVIEPIIMTRDHTKMLQGFQGANITVETDADGVRTIRLEGRGKLTQGV 240

QY 240 DVPGDPSSTAPFLVAALLVPGSDVTILNVLNMPRTGLTLTQEMGADIEVINPRLAG 299

Db 241 AVPGDPSAGFLVAALLVPGSDVTILNVLNMPRTGLTLTQEMGADIEVINPRLAG 300

QY 300 DVADLRVRSSTLKGVTVPEDRAPSMIDEPYILAVAAAFAGATVMNGLEELRVKESDR 359

Db 301 DVADLRVRSSTLKGVTVPEDRAPSMIDEPYILAVAAAFAGATVMNGLEELRVKESDR 360

QY 360 AVANGKLVGDCDEGETSLVVRGPDGKGLG---NASGAAVATHLDHRIAMSLVNGLV 416

Db 361 AVANGKLVGDCDEGETSLVVRGPDGKGLG---NASGAAVATHLDHRIAMSLVNGLV 420

QY 417 SENPVTDDATMIATSFPEFMDLMAGLAKI 447

Db 421 TEKPTIDDOAMIAISFPEFMDLMAGLAKI 451

RESULT 4

QY 421 Q9A2H2 PRELIMINARY; PRT; 443 AA.

AC Q9A2H2

DT 01-JUN-2001 (TremBLrel. 17, Created)

DT 01-JUN-2001 (TremBLrel. 17, Last sequence update)

DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)

DE 3-PROSPOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE.

GN CC3589.

OS Caulobacter crescentus.

OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;

OC Caulobacter.

OX NCBI_TaxID=69394;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 19089 / CB15;

RX MEDLINE=21173698; PubMed=11259647;

RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E., Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R., Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B., DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H., Kolonay J.F., Smith J., Craven M.B., Khouri H., Shetty J., Berry K., Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O., Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;

RT "Complete genome sequence of *Caulobacter crescentus*."

RT Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).

RL EMBL; AE006017; AAK25551.1; -

DR TIGR; CC3589; -

DR InterPro; IPR001986; EPPS_synthase.

DR Pfam; PF0275; EPPS_synthase; 1.

DR ProDom; PD001867; EPPS_synthase; 1.

DR PROSITE; PS00104; EPPS_SYNTHASE_1; UNKNOWN_1.

KW Transferase; Complete proteome.

SQ SEQUENCE 443 AA; 46075 MW; 45E8A1463E10B6EC CRC64;

Query Match 49.6%; Score 1135; DB 16; Length 443;

Best Local Similarity 55.8%; Pred. No. 4.5e-61;

Matches 251; Conservative 45; Mismatches 138; Indels 16; Gaps 6;

QY 2 LHGASRRPATAKSGSLGTVRIPGDKSISHRSPFPGGLASGETRITGLLEGEDVINTGK 61

Db 3 LAGLSAPGGA-----LRGIVRAPGDKSISHRSPFPGGLASGETRITGLLEGEDVINTGK 57

QY 62 AMOAGARLKEG-DTWIIDGVNGGLLAPLDFGNATGRLTMGLVGVYDFDSTFI 120

Db 58 AMOAGARLKEG-DTWIIDGVNGGLLAPLDFGNATGRLTMGLVGVYDFDSTFI 115

QY 121 GDASLTKRPMGRVLEPLRQMGVQLKATPGDRMPITLHGPKHAAPIYRVPMASAOVKS 180

Db 116 GDQSLGRPMGRVLEPLRQMGVQLKATPGDRMPITLHGPKHAAPIYRVPMASAOVKS 174

QY 181 VLLAGLNTPGITTVIEPIIMTRDHTKMLQGFQGANITVETDADGVRTIRLEGRGKLTQGV 237

Db 175 VLLAGLNTPGITTVIEPIIMTRDHTKMLQGFQGANITVETDADGVRTIRLEGRGKLTQGV 234

QY 238 VDVPCDPSSTAPFLVAALLVPGSDVTILNVLNMPRTGLTLTQEMGADIEVINPRLAG 297

Db 235 HVAVPGDPSAGFLVAALLVPGSDVTILNVLNMPRTGLTLTQEMGADIEVINPRLAG 294

QY 298 GEDVADLRVRSSTLKGVTVPEDRAPSMIDEPYILAVAAAFAGATVMNGLEELRVKESDR 357

Db 295 GEEVGDIATARYSOLKGVVPPERAPSMIDEPYILAVAAAFAGATVMNGLEELRVKESDR 354

QY 358 LSAVANGKLVGDCDEGETSLVVRGPDGKGLG---NASGAAVATHLDHRIAMSLVNGLV 417

Db 355 ISLTANGKLVGDCDEGETSLVVRGPDGKGLG---NASGAAVATHLDHRIAMSLVNGLV 410

QY 418 ENPVTDDATMIATSFPEFMDLMAGLAKI 447

Db 411 QAEVAVDEPGMIATSFPEFMDLMAGLAKI 440

RESULT 5

QY 418 Q9H269 PRELIMINARY; PRT; 746 AA.

AC Q9H269

DT 01-MAR-2001 (TremBLrel. 16, Created)

DT 01-MAR-2001 (TremBLrel. 16, Last sequence update)

DT 01-OCT-2001 (TremBLrel. 18, Last annotation update)

DE STILL FRAMESHIFT 3-PROSPOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE

DE PREPHENATE DEHYDROGENASE.

GN PA3164.

OS Pseudomonas aeruginosa.

OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;

OC Pseudomonas.

OX NCBI_TaxID=287;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 15692 / PA01;

RX MEDLINE=20437337; PubMed=10984043;

RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P., Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M., Garber R.L., Gollery L., Tolentino E., Westbrock-Wadman S., Yuan Y., Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M., Smith K.H., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T., Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;

RT "Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen."

RT Nature 406:959-964(2000).

RL EMBL; AE004740; RAG06552.1; -

DR InterPro; IPR001986; EPPS_synthase.

DR InterPro; IPR000205; NAD_binding.

DR Pfam; PF00275; EPPS_synthase; 1.

DR Pfam; PF02153; PDH; 1.

DR ProDom; PD001867; EPPS_synthase; 1.

DR PROSITE; PS00104; EPPS_SYNTHASE_1; UNKNOWN_1.

DR PROSITE; PS00885; EPPS_SYNTHASE_2; UNKNOWN_1.

KW Complete proteome.

SQ SEQUENCE 746 AA; 79320 MW; C2974B4BBF539E3A CRC64;

Query Match 43.2%; Score 989.5; DB 16; Length 746;

Best Local Similarity 48.5%; Pred. No. 6e-52;

Matches 214; Conservative 62; Mismatches 152; Indels 13; Gaps 3;

QY 12 ARKSSGSLGTVRIPGDKSISHRSPFPGGLASGETRITGLLEGEDVINTGKAMGARIR 71

Db 317 APOGSSGSLGTVRIPGDKSISHRSPFPGGLASGETRITGLLEGEDVINTGKAMGARIR 376

QY 72 -KEGDTWIIDGVNGGLLAPLDFGNATGRLTMGLVGVYDFDSTFI 130

Db	377	GPQNGRVTVHGVGLHGLKAPCPYILGNSGTSMRLLSGLLAAQPFDTLTLDASLSKRP	436
Qy	131	GRVLNPLREMGVQKSGDGLPVTLRGPKTPTTITVRPVMASQVKSASVALLAGLNTPGI	190
Db	437	NRVAKPLREMGAVIETGPEGRPTTIRGQRLTGMHYDMPMASQVKSCLLGLAGLYAAGE	496
Qy	191	TTVIEPIWTRDHTKMLQGFGANLTVETDADGVRTIRLEGRGKLTGGQVIDVPGDPSSTAF	250
Db	497	TSVTEPAPTRDHTERMLRGFGYPPVVEGS-----TAKVESGHKLSATHIEVPADISSAAF	551
Qy	251	PLVAALLVPGSDVTLNVLMPNTRTGLTLTLOEMGADIEVINPRLAGDEDVADLRVRSST	310
Db	552	PLVASIAEGSELVQLHVGINPTRVGVTEILRLMGDDLSENQREVGEPVADIRVRSAR	611
Qy	311	LKGVTVVPEDRAPSMIDIEYPIILAVAAFAEGATVMNGLEELRVKESDRLSAVANGKLNGV	370
Db	612	LKGIIDIPEDLVLAIDEPVLFVAAACAEGRVTLRGAELRVKESDRISQVNMADGLKALGV	671
Qy	371	DCDEGETSLVVRGPRDGGKGLGNASCAAVATHLDRHIAISFLVGLVSNPVTVDATMIA	430
Db	672	KAEPTPDGIVIEG-----GAGGGEVWAGHDHRIAMSFVSASLRASGPIRIHDCANVA	724
Qy	431	TSFPPEFMDLMAGLGAKEIUSD 451	
Db	725	TSFPNLFALCAQTGIRVAVEN 745	
RESULT	6		
ID	Q9PB21	PRELIMINARY; PRT; 454 AA.	
AC	Q9PB21		
DT	01-OCT-2000 (TRENBLrel. 15, Created)		
DT	01-OCT-2000 (TRENBLrel. 15, Last sequence update)		
DT	01-OCT-2001 (TRENBLrel. 19, Last annotation update)		
DE	3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE.		
GN	XF2324.		
OS	Xylella fastidiosa.		
OC	Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;		
OC	Xylella.		
OX	NCBI_TaxID=2371;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=9A5C;		
RX	MEDLINE=20365717; PubMed=10910347;		
RA	Simpson A.J.G., Reinach F.C., Artuda P., Abreu F.A., Acencio M.,		
RA	Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,		
RA	Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,		
RA	Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrier H.,		
RA	Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,		
RA	Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,		
RA	Faccinani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,		
RA	Fraga J.S., Franca S.C., Franco M.C., Frohne M., Furlan L.R.,		
RA	Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,		
RA	Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,		
RA	Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,		
RA	Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,		
RA	Marques M.V., Martins A.B.N., Madeira H.M.F., Marino C.L.,		
RA	Marques M.V., Martins A.B.N., Martins E.M.F., Matsukuma A.Y.,		
RA	Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,		
RA	Meon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,		
RA	Nhani A., Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,		
RA	de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,		
RA	Peixoto B.R., Perella G.A.G., Pereira H.A., Jr., Pesquero J.B.,		
RA	Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,		
RA	de Rosa V.E., Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,		
RA	da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A., Jr.,		
RA	da Silva J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,		
RA	de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshako M.H.,		
RA	Vallada H., Van Sluys M.A., Verjovsky-Almeida S., Vettore A.L.,		
RA	Zawo M.A., Zatz M., Meidanis J., Setubal J.C.;		
RT	"The genome sequence of the plant pathogen Xylella fastidiosa."		
RL	Nature 406:151-159(2000).		
DB	EMBL; AF040403; AF085123.1;		

DR	InterPro: IPR001986; EPSP_synthase.		
DR	Pfam: PF00375; EPSP_synthase; 1.		
DR	ProDom: PD001867; EPSP_synthase; 1.		
DR	PROSITE: PS00104; EPSP_SYNTHASE_1; UNKNOWN_1.		
DR	PROSITE: PS00885; EPSP_SYNTHASE_2; UNKNOWN_1.		
KW	Complete proteome.		
SQ	SEQUENCE 454 AA; 48266 MW; 45CCF074E6C0BA57 CRC64;		
Query Match	39.9%; Score 912.5; DB 16; Length 454;		
Best Local Similarity	46.9%; Pred. No. 1.4e-47;		
Matches	206; Conservative 60; Mismatches 158; Indels 15; Gaps 4;		
Qy	12	ARKSSGLSGTVRIPGDKSISHRSEFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIR	71
Db	23	AHQGTPLHGLVSLIPGDKSISHRVMAALADGTSRIDGFLAEADCTCAETLARLGVRIE	82
Qy	72	KEGDTW-IIDGVNGGGLLAPLDFGNATGCRLTGLVGVYDFDSTFTGDASLTKRPM	130
Db	83	TPLSTQRIHVHGVGVDGLQASHIPLDCGNAGTGMRLLAGLLVAQPFDSVLVDASLSKRP	142
Qy	131	GRVLNPLREMGVQKSGDGLPVTLRGPKTPTTITVRPVMASQVKSASVALLAGLNTPGI	190
Db	143	RRVTDPLSQMGARIDTSDGTPPLRIYGGOLLHGIDFISPVASAIKSAVLLAGLYARNE	202
Qy	191	TTVIEPIWTRDHTKMLQGFGANLTVETDADGVRTIRLEGRGKLTGGQVIDVPGDPSSTAF	250
Db	203	TVREPHPTRDYTERMLTAFGVIDIVSTGC-----ARLGGQRLCATDITIPADFSAAF	257
Qy	251	PLVAALLVPGSDVTLNVLMPNTRTGLTLTLOEMGADIEVINPRLAGDEDVADLRVRSST	310
Db	258	YLVAASVPGSDITLRAVGLNPRIGLITVLRMGANIVESNRHEOGGEPVVDLRVYAP	317
Qy	311	LKGVTVVPEDRAPSMIDIEYPIILAVAAFAEGATVMNGLEELRVKESDRLSAVANGKLNGV	370
Db	318	LOGTRVPEDLVADIMDEFPALFVAAAAGQTVVSGAAELRVKESDRLLAAVMTGLRLVG	377
Qy	371	DCDEGETSLVVRGPRDGGKGLGNASCAAVATHLDRHIAISFLVGLVSNPVTVDATMIA	430
Db	378	QVDETFADGATGGPIGHGTINSHG-----DHRIAMAFSIAGLSYSTVRIEDVANVA	430
Qy	431	TSFPPEFMDL--MAGLGAKI 447	
Db	431	TSFPDYETLARSAGFGLEV 449	
RESULT	7		
ID	Q9KCA6	PRELIMINARY; PRT; 431 AA.	
AC	Q9KCA6		
DT	01-OCT-2000 (TRENBLrel. 15, Created)		
DT	01-OCT-2000 (TRENBLrel. 15, Last sequence update)		
DT	01-OCT-2001 (TRENBLrel. 18, Last annotation update)		
DE	3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE (EC 2.5.1.19) (5-ENOLPYRUVYLSHIKIMATE-3-PHOSPHATE SYNTHASE) (EPSP SYNTHASE)		
DE	(EPSPS).		
GN	AROF OR BH1667		
OS	Bacillus halodurans.		
OC	Bacteria; Firmicutes; Bacillus/Clostridium group;		
OC	Bacillus/Staphylococcus group; Bacillus.		
OX	NCBI_TaxID=86665;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C-125 / JCM 9153;		
RX	MEDLINE=20512582; PubMed=11058132;		
RA	Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,		
RA	Fuji F., Hirana C., Nakamura Y., Ogasawara N., Kuhara S.,		
RA	Horikoshi K.;		
RT	"Complete genome sequence of the alkaliphilic bacterium Bacillus		
RT	halodurans and genomic sequence comparison with Bacillus subtilis."		
RL	Nucleic Acids Res. 28:4317-4331(2000).		
CC	-1- CATALYTIC ACTIVITY: PHOSPHOENOLPYRUVATE + 3-PHOSPHOSHIKIMATE =		
CC	ORTHOPOSPHATE + O(5)-(1-CARBOXYVINYL)-3-PHOSPHOSHIKIMATE.		

CC -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS FROM CHORISMATE OF THE
CC AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC -!- SIMILARITY: BELONGS TO THE EFSP SYNTHASE FAMILY.
DR EMBL; AP001512; BAB05386.1; -.
DR InterPro; IPR001986; EFSP_synthase.
DR Pfam; PF00275; EFSP_synthase; 1.
DR ProDom; PD001867; EFSP_synthase; 1.
DR PROSITE; PS00104; EFSP_SYNTHASE_1; 1.
DR PROSITE; PS00885; EFSP_SYNTHASE_2; 1.
DR PROSITE; PS00885; EFSP_SYNTHASE_2; 1.
KW Aromatic amino acid biosynthesis; Complete proteome; Transferase.
SQ SEQUENCE 431 AA; 45485 MW; 12F4FBE7BA0743D CRC64;

Query Match 38.6%; Score 884; DB 16; Length 431;
Best Local Similarity 44.2%; Pred. No. 6.9e-46;
Matches 188; Conservative 72; Mismatches 153; Indels 12; Gaps 5;
QY 15 SSGLSGTVIRPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEG 74
10 AKGLGKTIKVPDGSISHRAVMFEGALAKGTTTVEGFLPGADCLSTISCFQKLGVSIEQAE 69
75 DTWIIDGVNGGLLAPEAPLDFGNAATGCLTMGLVGVYDFDSTFIGDASLTKRPMGRVL 134
70 ERVTVKGGWGLRPSDILGVNSGTTTLTGLTSLTPFHSVITGDESIGKRPKRVY 129
135 NPLRMGVQVKSD-GDRLPVTLRGPKTPTPTIRYVPMAOVSAAVLLAGLNTPGITTV 193
130 EPLKSGAQIDGRDGNLTPLSIRGQL-KGIDFHPVSAOAKSAILLAGURAGKTSV 188
194 IEPIMTRDHTKMLQGFANLTVETDADGVRTIRLEGKGLTGQVIVDPGDSSTAFPLV 253
189 TEPAKTRDHTERMLEAFGNI---EKDGL-TVSIEGGQMLTGQVIVDPGDISSAAFLV 243
254 AALLVPGSDVTLNVLMPTRTGLTLTLOEMGADIEVINPRLAGGEDVADLRVRSSTLKG 313
244 AGAMVPHSRITLTNVGINPTFRAGILEVLKQMGATLAMENRVQGGEPVADLTETISVLOG 303
314 VTVPEDRAPSMIDEPYILAVAAFAEGATVMNGLEELRVKESDRLSAVANGKLNGVDCD 373
304 VEIGDIIIPRLIDELPIIATVATQASGRVTKDAELKVKETNRIDTVVSELTKLGAETH 363
374 EGETSLVVRGPDGKGLGNASGAATVATHLDRHIAFSLVNGVLSNPVTVDDATMIATSF 433
364 ATDDGMIEGPTPLAG----GVTVSSHGDHRIGMAMAIAALLAEKPVTEGTEAIVSY 418
434 PEFMD 438
419 PSFFD 423

RESULT 8
Q99283 ID Q99283 PRELIMINARY; PRT; 430 AA.
AC Q99283;
DT 01-JUN-2001 (TremBLrel. 17, Created)
DT 01-JUN-2001 (TremBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)
DE PUTATIVE 3-PHOSPHOSHIKIMATE 1-CARBOXYINYLTRANSFERASE
DE (EC 2.5.1.19).
GN AROA OR SPY1352.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SF370 / ATCC 700294 / SEROTYPE M1;
RX MEDLINE=21192684; PubMed=11296296;
RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Seaton S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";

RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663 (2001).
DR EMBL; AE006573; AAK34180.1; -.
DR InterPro; IPR001986; EFSP_synthase.
DR Pfam; PF00275; EFSP_synthase; 1.
DR ProDom; PD001867; EFSP_synthase; 1.
DR PROSITE; PS00104; EFSP_SYNTHASE_1; 1.
DR PROSITE; PS00885; EFSP_SYNTHASE_2; 1.
DR PROSITE; PS00885; EFSP_SYNTHASE_2; 1.
KW Transferase; Complete proteome.
SQ SEQUENCE 430 AA; 46692 MW; 3273C6B39020FB61 CRC64;
Query Match 35.8%; Score 820; DB 16; Length 430;
Best Local Similarity 43.8%; Pred. No. 5.1e-42;
Matches 189; Conservative 76; Mismatches 141; Indels 26; Gaps 9;
QY 13 RKSSG-LSGTVRIPDGSISHRSFMFGGLASGETRITGLLEGEDVINTKAMQAMGARIR 71
7 RTNAGPLQCTIQVPGDKSISHRAVLGAVAKGTRVKGGLKGLDVLSTQAFNGLGVRIE 66
72 KEGDTWIIDGVNGGLLAPEAPLDFGNAATGCLTMGLVGVYDFDSTFIGDASLTKRPMG 131
67 EKDDQLVIEGQFGQGLNAPCQTLNMGNSGTSMRLIAGLLAGQPFVSKMIGDESLSKRPMD 126
132 RVLNPLREMGVQVKS-PDGRPLVTLRGPKTPTPTIRYVPMAOVSAAVLLAGLNTPGI 190
127 RIVYPLKQMGVEISGETDRQFPPLQOGNRLQPIYITLPISSAQVKSAILLAALQAKGT 186
191 TVIPIPTRDHTKMLQGFANLTVETDADGVRTIRLEGKGLTGQVIVDPGDPSPSTAF 250
187 TQVEKEIRNTEEMIQFGGLIV---DGKR-ITLVGPQQLTAQETVPGDISSAAF 241
251 PLVAALLVPGSDVTLNVLMPTRTGLTLTLOEMGADI--EVINPRLAGGEDVADLRVRS 308
242 WLVAAGLIIPGSELLKNGVNPRTGILEVEMKGAQIVYEDMNKK----EQVTSIRVYV 297
309 STLKGVTPEDRAPSMIDEPYILAVAAFAEGATVMNGLEELRVKESDRLSAVANGKLN 368
298 SNMKGTTISGGLIIPRLIDELPIIATVATQAGTCTCKDAQELVETKDRIQVTVTDLNSM 357
369 GYVDCDEGETSLVVRGPDG---KGLGNASGAATVATHLDRHIAFSLVNGVLSNPVTV 424
358 GAN-----IKATADGMIIKGPVLYGANTSTYGDHRIGMTAIAALLVKQGQVHLD 408
425 DATMTATSPPEP 436
409 KEEAINTSYPTF 420
RESULT 9
Q92A85 ID Q92A85 PRELIMINARY; PRT; 428 AA.
AC Q92A85;
DT 01-DEC-2001 (TremBLrel. 19, Created)
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)
DE AROE PROTEIN
GN AROE OR LIN2037.
OS Listeria innocua.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Listeria.
OX NCBI_TaxID=1642;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CLIP 11262 / SEROVAR 6A;
RX PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Deloux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fshih H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kuhn M., Kunst F., Kurapat G.,
RA Madieno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,

DR	Nordstiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,	DR	PROSITE; PS00885; EPSP_SYNTHASE_2; UNKNOWN_1.
RA	Rommel B., Rose M., Schlueder T., Simoes N., Tierrez A.,	KW	Transferase.
RA	Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.,	SQ	SEQUENCE 428 AA; 45715 MW; 9E0F4FE4A893CA95 CRC64;
RT	"Comparative genomics of <i>Listeria species</i> ."		
RL	Science 294:849-852(2001).		
DR	EMBL; AL596170; CAC97267.1; ..		
DR	Listlist; LIN02037; ..		
KW	Complete proteome.		
SQ	SEQUENCE 428 AA; 45994 MW; 157B48C091A68FEB CRC64;		
Query Match			
Best Local Similarity 35.6%; Score 815.5; DB 16; Length 428;			
Matches 173; Conservative 83; Mismatches 155; Indels 13; Gaps 5;			
QY	17 GLSGTVIRIPGDKSHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKKGDT 76	QY	18 LSGTVIRIPGDKSHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKKGDTW 77
DB	9 GLVGEITVPDKSMHSRISIMFGAIAEGKTVIRHFLRADDCGLTGKAFKALGVKIEETDEE 68	DB	12 LOGTLVPSDKSHRSIMFGAIASSGKTTITNLFGRGDCLSLTAAPRSLGVNIEDDGT 71
QY	77 WIIDGVNGGLAPLAPDFGNAATGCLRTMGLVGVYDFDSTFGDASLTAKRPMGRVLP 136	QY	78 IIDGVNGGLAPLAPDFGNAATGCLRTMGLVGVYDFDSTFGDASLTAKRPMGRVLP 137
DB	69 IIVHGTGSDGLKQAGPGLDIGNSGTTIRLMGILAGRDFTVILGDSIAKRPNNRMLP 128	DB	72 TVEGRGFAGLKAKRTIDVNGSGTTIRLMGILAGCPFETRLAGDASIAKRPNNRMLP 131
QY	137 IREMGVQVKSDEGDRLL-DVTLRGPKTPPTIYRVPMASAOVKSALLAGLNTPGITTVIE 195	QY	138 REMGVQVKS-EDGDRLPVTLRGPKTPPTIYRVPMASAOVKSALLAGLNTPGITTVIEP 196
DB	129 LOEMGAKHGRDGESEFAPISIGNSLKRMVHMPVASAOVKSALLAGLNTPGITTVIE 188	DB	132 NQMAECOGVOOTEFPPISIRGTONLPIDYTPVSAQVKSALLAGLNTPGITTVIEP 191
QY	196 PIMTRDHTKMLQGFANLTVDADGVRTIRLEGRGLTGQVDPGDSAPFLVAA 255	QY	197 IMTRDHTKMLQGFANLTVDADGVRTIRLEGRGLTGQVDPGDSAPFLVAA 256
DB	189 KEKTRDHTKMLQGFANLTVDADGVRTIRLEGRGLTGQVDPGDSAPFLVAA 243	DB	192 EKTRDHTKMLQGFANLTVDADGVRTIRLEGRGLTGQVDPGDSAPFLVAA 246
QY	256 LVPDGSVTLNLMNPTRTGLILTOEMGADIEVINPRLAGGEDVADLVRSSLTGKVT 315	QY	257 LVPDGSVTLNLMNPTRTGLILTOEMGADIEVINPRLAGGEDVADLVRSSLTGK 312
DB	244 LITPGEITLHVLNPTRTGIFDVVEQMGSLVVKDSRSTGKLAGTVVVKSELKGE 303	DB	247 VVPDSEILLKNVGLNQRTGILDVKNMGSGVTIN-----EDENHSGDLLVKTSLT 300
QY	316 VPDGRAPSMIDEPYILAVAAFAEGATVMNGLEELRVKESDRLSAVANGLKLVGDCB 375	QY	313 GVTVPEDRAPSMIDEPYILAVAAFAEGATVMNGLEELRVKESDRLSAVANGLKLVGDC 372
DB	304 IGGDIIPRLIDEPYILAVAAFAEGATVMNGLEELRVKESDRLSAVANGLKLVGDCB 363	DB	301 ATEIGGAIIPRLIDEPYILAVAAFAEGATVMNGLEELRVKESDRLSAVANGLKLVGDC 360
QY	376 ETSVVRGRPDGKGLGNASGAATVHLDRHRTAMSLVVG-LVSENPVTVDDATMIATSP 434	QY	373 DEGETSVVRGRPDGKGLGNASGAATVHLDRHRTAMSLVVG-LVSENPVTVDDATMIAT 431
DB	364 EDGLIRGTP-----LHAANTVSYGDRHGMQLQIAALLVEDGDVLDRAEAVSVSP 417	DB	361 TPTDGLIHGPT-----SLHGRKVTSYGDRHGMQLQIAALLVEDGDVLDRAEAVSV 414
QY	435 EFMD 438	QY	432 SFPEFMD 438
DB	418 TTFE 421	DB	415 SYPAFFD 421
RESULT 10			
QY	Q9ANY6 PRELIMINARY; PRT; 428 AA.	QY	Q99U25 PRELIMINARY; PRT; 432 AA.
AC	Q9ANY6; (1)	AC	Q99U25; (1)
DT	01-JUN-2001 (TREMELrel. 17, Created)	DT	01-JUN-2001 (TREMELrel. 17, Created)
DT	01-JUN-2001 (TREMELrel. 17, Last sequence update)	DT	01-JUN-2001 (TREMELrel. 17, Last sequence update)
DT	01-OCT-2001 (TREMELrel. 18, Last annotation update)	DT	01-DEC-2001 (TREMELrel. 19, Last annotation update)
DE	3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE	DE	3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE.
GN	ARO6 OR SA1297 OR SAV1464.	GN	ARO6 OR SA1297 OR SAV1464.
OS	Staphylococcus aureus (strain N315), and	OS	Staphylococcus aureus (strain N315), and
OC	Bacteria; Firmicutes; Bacillus/Clostridium group;	OC	Staphylococcus aureus (strain Mu50).
OC	Bacillus/Staphylococcus group; Staphylococcus.	OC	Bacteria; Firmicutes; Bacillus/Clostridium group;
OX	NCBI_TaxID=15878;	OX	Bacillus/Staphylococcus group; Staphylococcus.
RN	[1]	RN	NCBI_TaxID=15878; 158878;
RP	SEQUENCE FROM N.A.	RP	[1]
RC	Staphylococcus aureus (strain N315), and S.aureus (strain Mu50);	RC	SEQUENCE FROM N.A.
RC	SPECIES-S.aureus (strain N315), and S.aureus (strain Mu50);	RC	SPECIES-S.aureus (strain N315), and S.aureus (strain Mu50);
RX	MEDLINE=21311952; PubMed=11418146;	RX	MEDLINE=21311952; PubMed=11418146;
RA	Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,	RA	Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA	Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,	RA	Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA	Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,	RA	Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA	Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,	RA	Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA	Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,	RA	Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA	Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,	RA	Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA	Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.	RA	Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.
RT	"Whole genome sequencing of methicillin-resistant <i>Staphylococcus aureus</i> ."	RT	"Whole genome sequencing of methicillin-resistant <i>Staphylococcus aureus</i> ."
RL	Lancet 357:1225-1240(2001).	RL	Lancet 357:1225-1240(2001).
DR	EMBL; AP003134; BAB42557.1; ..	DR	EMBL; AP003134; BAB42557.1; ..
DR	EMBL; AP003362; BAB57626.1; ..	DR	EMBL; AP003362; BAB57626.1; ..
DR	InterPro; IPR001986; EPSP_synthase.	DR	InterPro; IPR001986; EPSP_synthase.

DR Pfam: PF00275; EPSP_synthase; 1.
DR ProDom: PD001867; EPSP_synthase; 1.
DR PROSITE: PS00104; EPSP_SYNTHASE_1; 1.
DR PROSITE: PS00885; EPSP_SYNTHASE_2; UNKNOWN_1.
KW Transferase; Complete proteome.
SQ SEQUENCE 432 AA; 47068 MW; A42102057AD15C72 CRC64;

Query Match 26.7%; Score 610; DB 16; Length 432;
Best Local Similarity 34.0%; Pred. No. 2.7e-29;
Matches 146; Conservative 80; Mismatches 181; Indels 22; Gaps 6;

QY 15 SSGLSGTVPKDKSISHSFSGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEG 74
DB 10 SGPLKEIEVPGDKSWTHRAIMLASLAEGTSNIYKPLLEGDCRRWMDIFRLGVDIKDE 69

QY 75 DTWIDGVNGGGLLAPEAPLDFGNAATGRLTMGLGVVYDFDSTFGDASLTKRPMGRVL 134
DB 70 DKLNVNSPGYKAFKTPHQVLYTNGSGTTRLLAGLLSGLGIESVLSGKRPMDRVL 129

QY 135 NPLREMGVQVKSDEGDRPLVTLRGPKTPITYRVPMSAQVSAVLLAGLNTPTGTTVI 194
DB 130 RPLKMDANIEGIDNYTPLIK-PSVIGKINQMEVASAQVSAVLLAGLNTPTGTTVIK 188

QY 195 EPIMTRDHTKMLQGF-----GANLTVEDADGVRTIRLEGRGKLTGVTDVPGDPSSTA 249
DB 189 ELDVSRNHTETFRFNIDIERISITTPDAIQIHKPAD-----FHVPGDISAA 240

QY 250 FPLVAALLVPGSDVTILNLMNPTRLGLTLQEMGADIEVINPRLAGGEDVADLRV-S 308
DB 241 FFIVAALITPESDVTIHNVGINPTSGIIDIIVKMGNGIQLFN-QTTGAEPASIRIQYT 299

QY 309 STLKGVTPEDRAPSMIDEPYILAVNAFAEGATVNGLEELRYKESDRLSAVANGLKN 368
DB 300 PMLQPTTIEGVLPAIDELPVALCTQAVGTSTIKDAELKVKETNRIDTADMNLNL 359

QY 369 GYDCEGETSLVVRGPDGKGLGNAGAAVATHLDRIAMSLFVGLVSENPTVVDATM 428
DB 360 GFELQPTNDGLIHHSE-----FKTNATVDSLTDRHGMMLAVASLLSEPKIKQFDA 413

QY 429 IATSPPEFM 437
DB 414 VNVSPFGFL 422

RESULT 12
Q9RHZ8 PRELIMINARY; PRT; 207 AA.
Q9RHZ8; MEDLINE=20504483; PubMed=11016950;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE 5-ENOLPYRUVLSHKIMATE 3-P SYNTHASE (FRAGMENT).
GN AROF.
OS Pseudomonas stutzeri (Pseudomonas perfectomarina).
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=316;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JM300;
RX MEDLINE=92013931; PubMed=1915906;
RA Fischer R.S., Zhao G., Jensen R.A.;
RT "Cloning, sequencing, and expression of the P-protein gene (phea) of
RT Pseudomonas stutzeri in Escherichia coli: implications for
RT evolutionary relationships in phenylalanine biosynthesis";
RL J. Gen. Microbiol. 137:1293-1301(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-JM300;
RX MEDLINE=99298294; PubMed=10368439;
RA Xie G., Bonner C.A., Jensen R.A.;
RT "A probable mixed-function supraoperon in Pseudomonas exhibits gene

RT organization features of both intergenomic conservation and gene
RT shuffling";
RL J. Mol. Evol. 49:108-121(1999).
DR EMBL: AF038578; AAD47363.1; -.
DR InterPro: IPR001986; EPSP_synthase.
DR Pfam: PF00275; EPSP_synthase; 1.
DR ProDom: PD001867; EPSP_synthase; 1.
DR PROSITE: PS00104; EPSP_SYNTHASE_1; UNKNOWN_1.
DR NON_TER 207
SQ SEQUENCE 207 AA; 21780 MW; 7231191C72A21D6B CRC64;

Query Match 20.6%; Score 470.5; DB 2; Length 207;
Best Local Similarity 51.5%; Pred. No. 2.7e-21;
Matches 103; Conservative 22; Mismatches 70; Indels 5; Gaps 2;

QY 12 ARKSSGLSGTVPKDKSISHSFSGGLASGETRITGLLEGEDVINTGKAMQAMGARIR 71
DB 10 ANPGSGSLTGOLRYPGDKSISHSRSMILGSLAEGTTEXEGFLEGEDALATXQAFRGMGVI- 68

QY 72 KEG---DTWIDGVNGGGLLAPEAPLDFGNAATGRLTMGLGVVYDFDSTFGDASLTKR 128
DB 69 -EGPHQGRVTVHGVGLHQAQPPGPIYLGNSGTSMRLLAGLLAAQFPDFTLSGDASLTKR 127

QY 129 PMGRVNLPLREMGVQVKSDEGDRPLVTLRGPKTPITYRVPMSAQVSAVLLAGLNTPT 188
DB 128 PMNRVAKPLREMGAVIETAEAGRPPLTIRGGKKLSGHHYDMPMASAQVSCLLLAGLYAA 187

QY 189 GITTVEIPTNTRDHTKMLQ 208
DB 188 GKTSVTEPAPTRDHTERMQ 207

RESULT 13
Q9HQC1 PRELIMINARY; PRT; 439 AA.
Q9HQC1; MEDLINE=20504483; PubMed=11016950;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE 3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE.
GN PSC OR VNG1232G.
OS Halobacterium sp. (strain NRC-1).
OC Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae;
OC Halobacterium
OX NCBI_TaxID=64091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Weiti R., Goo Y.A.,
RA Leitzauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.N., Dale H.,
RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;
RT "Genome sequence of Halobacterium species NRC-1";
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
DR EMBL: AE005049; AAG19594.1; -.
DR InterPro: IPR001986; EPSP_synthase.
DR Pfam: PF00275; EPSP_synthase; 1.
DR ProDom: PD001867; EPSP_synthase; 1.
DR PROSITE: PS00104; EPSP_SYNTHASE_1; UNKNOWN_1.
DR PROSITE: PS00885; EPSP_SYNTHASE_2; UNKNOWN_1.
KW Transferase; Complete proteome.
SQ SEQUENCE 439 AA; 44333 MW; 48CAD75A1FOECB89 CRC64;

Query Match 20.6%; Score 470.5; DB 17; Length 439;
Best Local Similarity 31.3%; Pred. No. 7.6e-21;
Matches 139; Conservative 76; Mismatches 190; Indels 39; Gaps 13;

Qy	16	SGLSGTVRIPGDKSIHRSFMEFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGD	75
Db	17	SVKRGARAPPSKYTHRALAAGYADGETVVRDPLVSADTRATARAVELLGGAAARENG	76
Qy	76	TWIDGVNGGALLAPEAPLDFGNAATCRLTMGLGVGVDFDSTFIGDASLTKRPMGRVLN	135
Db	77	DWVTVFGSRPAI-PDAVIDCANSCTMRLVLTAAALADGTTVLTGDESLEAPRPHGPLLD	135
Qy	136	PLREMGVQVKSEDDG-RLPVTLRGPKTPTITYRVP-MASAQVKSAVLLAGLNT-----P	188
Db	136	ALSGLGGTARSTRGNGQAPLVDPVSGSVA--LPGDVSSQFVTALIMAGAVTETGIET	193
Qy	189	GITTVIEPIMRDHTKMLQGFANLTVETDADGVRTIRLEGRG-----KLTGOVIDVPDG	244
Db	194	DLTTELKSAPYVDIPLVDLDFAGVAS-ET-AAGTRV-----RGQAYAPSGAEYAVPGD	246
Qy	245	PSSTAFFLVAALLVPGSDVTILNLMNPTRTG---LILTLOEMGADIIEVINPRLAGGEDV	301
Db	247	FSSASYLLAAGALAAADGAANVVVEGMHPSAQGDAAIVDVLERMGADID-----WDTES	299
Qy	302	ADLRVRSSTLKGTVVPEDRAPSMIDEYPIILAVAAFAEAGATVMNGLEELRVKESDLSAV	361
Db	300	GVTIVQSRSELSGVEGVADTDPOLL---PTIAVLGAADGTTTRITDAEHVRYKETDRVAAM	356
Qy	362	ANGLKLVGDCDEGETSLVVRGRPDGKGLGNASGAATHLDHRIAMSFVLMGLVSENPV	421
Db	357	BSLSKLGASVEERDELVVRG-----GDTLSGASVDGRGDRHRLVMAVAGLVADGET	411
Qy	422	TVDDATMIATSPFPEMDLMAGLGA 445	
Db	412	TIAGSEHVDVSPDFFEVLGLGA 435	
RESULT	14		
Q9L213		PRELIMINARY; PRT; 440 AA.	
AC	Q9L213		
DT	01-OCT-2000 (TremBLrel. 15, Created)		
DT	01-OCT-2000 (TremBLrel. 15, Last sequence update)		
DT	01-JUN-2001 (TremBLrel. 17, Last annotation update)		
DE	3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE.		
GN	ARO.		
OS	Streptomyces coelicolor.		
OC	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;		
OC	Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.		
OX	NCBI_TaxID=1902;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RL	STRAIN=A3(2);		
RL	Brown S.P., Harris D.;		
RL	Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=A3(2);		
RA	Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;		
RL	Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=A3(2);		
RX	MEDLINE=97000351; PubMed=8843436;		
RA	Redenbach M., Kieser H.M., Denapaita D., Eichner A., Cullum J.,		
RA	Kinashi H., Hopwood D.A.;		
RT	"A set of ordered cosmids and a detailed genetic and physical map for		
RT	the 8 Mb Streptomyces coelicolor A3(2) chromosome.";		
RL	Mol. Microbiol. 21:77-96(1996).		
DR	EMBL; ALI38598; CAB71266.1; -		
DR	InterPro; IPR001986; EPSP_synthase.		
DR	Protein; IPR000408; RCP1.		
DR	Pfam; PF00275; EPSP_synthase; 1.		
DR	ProDom; PD001867; EPSP_synthase; 1.		
DR	ProSITE; PS00885; EPSP_SYNTHASE.2; 1.		
DR	ProSITE; PS00626; RCP1_2; UNKNOWN_1.		
KW	Transferase.		
Qy	11	TARKSSGLSGTVRIPGDKSIHRSFMEFGGLASGETRITGLLEGEDVINTGKAMQAMG---	67
Db	15	TIRLTGDETFEVLGSKSYNRYLAISLGOETVIDNALLSDDTVFSRAIETFGHVT	74
Qy	68	-----ARIRKEGDTWIIDGVNGGALLAPEAPLDFGNAATCRLTMGLGVGVDFDSTFI	120
Db	75	CDIDHATARIR-----VTPTGR-PMRAPSIEDIFVGAGCTPLRFLISMAGHADGTTIIT	126
Qy	121	GDASLTTRPMGRVLNPLREMGVQVKSEDDG-RLPVTLRGPKTPTITYRVPMAQAQVKA	179
Db	127	GNARQERPMGDLKALPALGVDAATVATRGNGSPVVRVGGSGKGGATSGISGAVSSQFTSS	186
Qy	180	VLLAGLNTPGIT--TVIEPIMRDHTKMLQGFANLTVETDADGVRTIRLEGRKLTGQ	237
Db	187	LIINALRAQTDTEITISDDLVSQKPYVENTLAGL-AEMGVSVDRDGYRFTVPSCQARGG	245
Qy	238	VIDVPGDPSSTAFPLVAA-----LLVPGSDVTILNLMNPTRTGTLILTLOEMGADIEVI	291
Db	246	QVTVEPDASGMSYFLAAAILQSRVVIIPG-----IGAGSHOGDVLVQALERMGRTEV-	299
Qy	292	NPLRAGGEDVADLRVRSSTLKGTVVPEDRAPSMIDEYPIILAVAAFAEAGATVMNGLEELR	351
Db	300	-----GDD--SIIVTGGPLRGIDIDEMPDVV---PSLAIVAAYAAGCTTRITNIASLR	348
Qy	352	VKESDRLSAVANGKLKLVGDCDEGETSLVVRGRPDGKGLGNASGAATHLDHRIAMSFVLM	411
Db	349	VKECDRIAAVTELRKMGIDVEEHSADAMYITG-----GTPHGAVIDTYDDHRIAMTFA	401
Qy	412	VNGLVSENPVTVDDATMIATSPFPEMDLMAGLGA 448	
Db	402	IGGLRTEG-VWIKDPCGVAKSPFAFWQTLDTLHPDLE 437	
RESULT	15		
Q9KR80		PRELIMINARY; PRT; 426 AA.	
ID	Q9KR80		
AC	Q9KR80;		
DT	01-OCT-2000 (TremBLrel. 15, Created)		
DT	01-OCT-2000 (TremBLrel. 15, Last sequence update)		
DT	01-DEC-2001 (TremBLrel. 19, Last annotation update)		
DE	3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE (EC 2.5.1.19) (5-		
DE	ENOLPYRUVYLSHIKIMATE-3-PHOSPHATE SYNTHASE) (EPSP SYNTHASE)		
DE	(EPSPS).		
GN	VC1732.		
OS	Vibrio cholerae.		
OC	Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.		
OX	NCBI_TaxID=666;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=EL TOR N16961 / SEROTYPE O1;		
RX	MEDLINE=20406833; PubMed=10952301;		
RA	Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,		
RA	Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Mayam L.A.,		
RA	Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,		
RA	Ermlaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,		
RA	McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,		
RA	Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,		
RA	Fraser C.M.;		
RT	"DNA sequence of both chromosomes of the cholera pathogen Vibrio		
RT	cholerae.";		
RL	Nature 406:477-483(2000).		
CC	-!- CATALYTIC ACTIVITY: PHOSPHOENOLPYRUVATE + 3-PHOSPHOSHIKIMATE -		
CC	ORTHOPHOSPHATE + O(5)-(1-CARBOXYVINYL)-3-PHOSPHOSHIKIMATE.		
CC	-!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS FROM CHORISMATE OF THE		
CC	AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).		
CC	-!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.		

Search completed: August 15, 2002, 14:05:44
Job time: 413 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
 Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 15, 2002, 13:58:47 ; Search time 56.25 Seconds
 (without alignments)
 898.464 Million cell updates/sec

Title: US-09-464-099A-70
 Perfect score: 2288
 Sequence: 1 MLHGASRPATARKSSGLSG.....FMDLMAGLAKIELSDTKAA 455

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : A_Genesec-052802.*

1:	/SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1980.DAT.*
2:	/SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1981.DAT.*
3:	/SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1982.DAT.*
4:	/SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1983.DAT.*
5:	/SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1984.DAT.*
6:	/SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1985.DAT.*
7:	/SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1986.DAT.*
8:	/SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1987.DAT.*
9:	/SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1988.DAT.*
10:	/SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1989.DAT.*
11:	/SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1990.DAT.*
12:	/SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1991.DAT.*
13:	/SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1992.DAT.*
14:	/SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1993.DAT.*
15:	/SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1994.DAT.*
16:	/SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1995.DAT.*
17:	/SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1996.DAT.*
18:	/SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1997.DAT.*
19:	/SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1998.DAT.*
20:	/SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1999.DAT.*
21:	/SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA2000.DAT.*
22:	/SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2288	100.0	455	13 AAR26449	CP4-EPSPS. Synthe
2	2288	100.0	455	19 AAW34685	CP4-EPSPS protein.
3	2288	100.0	527	22 AAM52214	EPSPS SEQ ID NO 3.
4	2282	99.7	455	13 AAR22300	Class II EPSPS enz
5	2282	99.7	455	18 AAW34683	Class II EPSP synt
6	2282	99.7	455	18 AAW24474	Class II EPSPS for
7	2282	99.7	455	19 AAW71609	Agrobacterium sp.
8	2282	99.7	455	22 AAE05053	Agrobacterium sp.
9	1900.5	83.1	449	13 AAR22301	Class II EPSPS enz
10	1900.5	83.1	449	13 AAR22302	Class II EPSPS enz
11	1900.5	83.1	449	18 AAW34684	Class II EPSP synt

12	1900.5	83.1	449	18 AAW34685	Class II EPSP synt
13	1900.5	83.1	449	18 AAW24479	Class II EPSPS use
14	1900.5	83.1	449	18 AAW24480	Class II EPSPS use
15	1900.5	83.1	449	19 AAW71611	Pseudomonas sp. st
16	1900.5	83.1	449	19 AAW71610	Achromobacter sp.
17	1900.5	83.1	449	22 AAE05054	Achromobacter sp.
18	1900.5	83.1	449	22 AAE05055	Pseudomonas sp. st
19	867.5	37.9	447	18 AAW34688	Class II EPSP synt
20	867.5	37.9	447	18 AAW24487	Class II EPSPS for
21	867.5	37.9	447	19 AAW71619	Synechocystis sp.
22	867.5	37.9	447	22 AAE05072	Synechocystis sp.
23	806	35.2	443	18 AAW34689	Class II EPSP synt
24	806	35.2	443	18 AAW24488	Class II EPSPS for
25	806	35.2	443	19 AAW71620	Dichelobacter nodo
26	806	35.2	443	22 AAE05073	Dichelobacter nodo
27	803.5	35.1	427	22 AAB48177	S. pneumoniae aroA
28	803.5	35.1	427	22 AAB37094	Streptococcus pneu
29	803.5	35.1	427	22 AAB48500	Streptococcus pneu
30	798.5	34.9	427	20 AAW97388	5-enolpyruvylshiki
31	798.5	34.9	427	21 AAY56504	Streptococcus pneu
32	798.5	34.9	427	22 AAB31158	A 5-enolpyruvylsh
33	785	34.3	428	18 AAW34686	Class II EPSP synt
34	785	34.3	428	18 AAW24481	Class II EPSPS use
35	785	34.3	428	19 AAW71617	Bacillus subtilis
36	785	34.3	428	22 AAE05070	Bacillus subtilis
37	704.5	30.8	415	20 AAW97389	protein encoded by
38	704.5	30.8	415	21 AAY56505	Streptococcus pneu
39	704.5	30.8	415	22 AAB31159	A 5-enolpyruvylsh
40	704.5	30.8	415	22 AAB48178	S. pneumoniae aroA
41	704.5	30.8	415	22 AAB37095	Streptococcus pneu
42	704.5	30.8	415	22 AAB48501	Streptococcus pneu
43	577	25.2	430	18 AAW34687	Class II EPSP synt
44	577	25.2	430	18 AAW24482	Class II EPSPS use
45	577	25.2	430	19 AAW71618	Staphylococcus aur

ALIGNMENTS

RESULT 1

AAR26449

ID AAR26449 standard; Protein; 455 AA.

AC AAR26449;

XX

DT 28-JAN-1993 (first entry)

XX

DE CP4-EPSPS.

XX

KW PMON11030; CTP2; CP4; EPSPS; chloroplast transit peptide;

KW 5-enolpyruvyl-3-shikimate phosphate synthetase; ACC; ethylene;

KW Pseudomonas chloroaphis 6G5.

XX

OS Synthetic.

XX

PN WO9212249-A.

XX

PD 23-JUL-1992.

XX

PF 17-DEC-1991; 91WO-US09437.

XX

PR 26-DEC-1990; 90US-0632440.

XX

XX (MONS) MONSANTO CO.

PA Kishore GM, Klee HJ;

PI WPI; 1992-284334/34.

XX N-PSDB; AAQ27201.

DR

XX Delaying fruit ripening and senescence in plants - by controlling

PT ethylene prodn., pref. by expression of

PT 1-amino:cyclopropane-1-carboxylic acid deaminase

XX PS Disclosure; Page 64-66; 110pp; English.

XX CC The sequences given in AAR26448 and AAR26449 are encoded by genes which

CC were used in the construction of an expression plasmid pMON11030 which

CC was used to transform petunia plants. This plasmid contained the genes

CC for chloroplast transit peptide (CTP2) (AAR26448) and the CP4 synthetic

CC 5-enolpyruvyl-3-shikimate phosphate synthetase (EPSPS) gene (AAR26449)

CC which is capable of conferring resistance to glyphosate. The plasmid

CC also contained the 1-aminocyclopropane-1-carboxylic acid (ACC)

CC deaminase gene from pseudomonas chlororaphis (see also AAR27199). The

CC transformed plants had ethylene levels reduced to about one half that

CC of the control, untransformed plants. It is expected that such plants

CC will show reduced senescence of flowers and leaves when compared to

CC untransformed plants.

XX CC

XX SQ Sequence 455 AA;

Query Match 100.0%; Score 2288; DB 13; Length 455;

Best Local Similarity 100.0%; Pred. No. 3.4e-181;

Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSPFGLASGETRITGLLEGEDVINTG 60

DB 1 mlhgassrpataarkssglsrgtvtvripgdksishrsfmfglasgetritgllegedvintg 60

QY 61 KAMQAMGARIRKEGDTWIIDGVNGGLLAPEAPLDFGNAATGCRLTMTGLVGVYDFDSTFI 120

DB 61 kamqamgarirkegdtwiidgvnggllapeapldfnaatgcrltmtglvgyvdfdstfi 120

QY 121 GDASLTKRPMGRVLPNPLREMGVQVKSDEGDRLPVTLRGPKTPTPTTYRVPMSAQAQVKSAAV 180

DB 121 gdasltkrpmgrvlnplremgvqvksedgdrilpvtlrgpktptpttyrvpmasaqvksaav 180

QY 181 LLAGLNTPGITTVIEIPIMTRDHTTEKMLQGFANLTVETDADGVRTIRLEGRKGLTQGVID 240

DB 181 llaglntpgittvievipimtrdhttekmqlqfganltvetdadvrtirlegrkgltgqv 240

QY 241 VPGDPSSTAPFLVAALLVPGSDVTILNVLNMPNTRTGLIILTLOEMGADIEVINPRLAGGED 300

DB 241 vpgdpsstaflvaallvpgsdvtlinvlnmpntrtgliltloemgadievlnprlagged 300

QY 301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPIILAVAAFAEGATVMNGLEELRVKESDRLSA 360

DB 301 vadlrvrssltkgtvtpedrapsmideypilavaafaegatvmngleelrvkesdrlsa 360

QY 361 VANGKLINGVDCDEGETSLVVRGPDGKGLGNASGAATVATHLDHRTAMSLVGLVSENP 420

DB 361 vangklngvdcdegetslvvrpgdgkglgnasgaavathldhrtamsflvmlvsenp 420

QY 421 VTVDATMIATSPFEMDLMAGLAKIELSDTKAA 455

DB 421 vtddatmiatspfemdlmaglakielsdtkaa 455

RESULT 2

AAR39426

ID AAR39426 standard; Protein; 455 AA.

XX AC

XX AA39426;

XX DT

XX 19-MAY-1998 (first entry)

XX DE

XX CP4-EPSPS protein.

XX KW 1-aminocyclopropane-1-carboxylic acid deaminase; ACC deaminase;

XX KW fruit ripening; ethylene production; glyphosate resistance;

XX KW 5-enolpyruvyl-3-shikimate phosphate synthase; EPSPS; CP4.

XX OS Synthetic.

XX PN U55702933-A.

XX PD 30-DEC-1997.

XX PF 06-NOV-1995; 95US-0553943.

XX PR 17-DEC-1991; 91US-0809457.

XX PR 26-DEC-1990; 90US-0632440.

XX PR 06-NOV-1995; 95US-0553943.

XX PA (MONS) MONSANTO CO.

XX PI Kishore GM, Klee HJ;

XX PT WPI: 1998-076419/07.

XX DR N-PSDB; AAV09719.

XX PT Production of plants with delayed ripening - using DNA encoding

XX PT 1-amino:cyclo:propane-1-carboxylic acid deaminase

XX PS Example 9; Fig 21; 56pp; English.

XX CC This sequence represents the synthetic CP4 5-enolpyruvyl-3-shikimate

CC phosphate synthase (EPSPS) gene which is capable of conferring

CC resistance to glyphosate and is used in a novel method for producing

CC fruit-bearing plants with delayed ripening. The method involves the

CC expression of a 1-aminocyclopropane-1-carboxylic acid (ACC) deaminase

CC gene in a plant at a level sufficient to reduce ethylene production in

CC the fruit.

XX SQ Sequence 455 AA;

Query Match 100.0%; Score 2288; DB 19; Length 455;

Best Local Similarity 100.0%; Pred. No. 3.4e-181;

Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSPFGLASGETRITGLLEGEDVINTG 60

DB 1 mlhgassrpataarkssglsrgtvtvripgdksishrsfmfglasgetritgllegedvintg 60

QY 61 KAMQAMGARIRKEGDTWIIDGVNGGLLAPEAPLDFGNAATGCRLTMTGLVGVYDFDSTFI 120

DB 61 kamqamgarirkegdtwiidgvnggllapeapldfnaatgcrltmtglvgyvdfdstfi 120

QY 121 GDASLTKRPMGRVLPNPLREMGVQVKSDEGDRLPVTLRGPKTPTPTTYRVPMSAQAQVKSAAV 180

DB 121 gdasltkrpmgrvlnplremgvqvksedgdrilpvtlrgpktptpttyrvpmasaqvksaav 180

QY 181 LLAGLNTPGITTVIEIPIMTRDHTTEKMLQGFANLTVETDADGVRTIRLEGRKGLTQGVID 240

DB 181 llaglntpgittvievipimtrdhttekmqlqfganltvetdadvrtirlegrkgltgqv 240

QY 241 VPGDPSSTAPFLVAALLVPGSDVTILNVLNMPNTRTGLIILTLOEMGADIEVINPRLAGGED 300

DB 241 vpgdpsstaflvaallvpgsdvtlinvlnmpntrtgliltloemgadievlnprlagged 300

QY 301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPIILAVAAFAEGATVMNGLEELRVKESDRLSA 360

DB 301 vadlrvrssltkgtvtpedrapsmideypilavaafaegatvmngleelrvkesdrlsa 360

QY 361 VANGKLINGVDCDEGETSLVVRGPDGKGLGNASGAATVATHLDHRTAMSLVGLVSENP 420

DB 361 vangklngvdcdegetslvvrpgdgkglgnasgaavathldhrtamsflvmlvsenp 420

QY 421 VTVDATMIATSPFEMDLMAGLAKIELSDTKAA 455

DB 421 vtddatmiatspfemdlmaglakielsdtkaa 455

RESULT 3

AAM52214

ID AAM52214 standard; Protein; 527 AA.

XX

AC AAM52214;
 XX 12-FEB-2002 (first entry)
 XX EPSPS SEQ ID NO 3.
 DE
 XX Transgenic plant; herbicide resistance; EPSPS; GOX; Petunia hybrida;
 KW 5-enolpyruvylshikimate-3-phosphate synthase; glyphosate oxidoreductase;
 KW protoporphyrinogen IX oxidase; immunoglobulin; Agrobacterium; strain CP4;
 KW Chloroplast transit peptide.
 XX
 OS Petunia hybrida.
 OS Agrobacterium sp.
 XX
 PN JP2001190168-A.
 XX
 PD 17-JUL-2001.
 XX
 XX 27-OCT-2000; 2000JP-0328811.
 XX
 XX 29-OCT-1999; 99JP-0310244.
 XX
 XX (SUMO) SUMITOMO CHEM CO LTD.
 XX
 XX WPI; 2001-605307/69.
 DR N-PSDB; ABA02854.
 XX
 XX New herbicide-resistant plant -
 PT
 XX
 PS Examples; Page 46-48; 66pp; Japanese.
 XX
 CC The invention relates to a transgenic plant which shows resistance to a
 CC herbicide in an amount inhibiting natural
 CC 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) activity of the plant
 CC and having at least one enzymatic activity selected from:
 CC (1) EPSPS activity different from natural EPSPS activity of the plant or
 CC (2) glyphosate oxidoreductase (GOX) activity different from the natural
 CC GOX activity of the plant; and in which a gene encoding a protein having
 CC the following properties: (a) combines specifically with a substance
 CC participating to the herbicidal activity a herbicide of
 CC protoporphyrinogen IX oxidase inhibiting type; (b) has substantially no
 CC denaturing activity on a substance to which said protein combines
 CC specifically; and (c) contains substantially no framework region of the
 CC variable region of immunoglobulin. The present sequence is that of the
 CC Petunia hybrida EPSPS chloroplast transit peptide and the Agrobacterium
 CC sp. strain CP4 EPSPS gene.
 XX
 XX Sequence 527 AA;

Query Match 100.0%; Score 2288; DB 22; Length 527;
 Best Local Similarity 100.0%; Pred. No. 4.2e-181;
 Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLHGASSRPATARKSSGLSGTVPRIKSGTSHRSFPGGLASGETRITGLLEGEDVINTG 60
 |||||
 Db 73 mlhgassrpataarkssglsgtvpripksgtshrsfmgfpglasgetritgllegedvintg 132
 |||||

QY 61 KAMQANGARIRKEGDTWIIDGVNGGLLAPEAPLDFGNAATGCRLLTMGLVGVYDFDSTFI 120
 |||||
 Db 133 kamqangarirkegdtwiidgvnggllapeapldfgnaatgcrlltmglvgvydfdstfi 192
 |||||

QY 121 GDASLTKRPMGRVNLPLREMGVQVKSEDDGRLPVTLRGPKTPPIYRVPMSAQVKSAY 180
 |||||
 Db 193 gdasltkrmpgrvnlplremgvqvkseddgdlpvtlrgpktptpityrvpmsaqvksav 252
 |||||

QY 181 LLAGLNTPGITTVIEPIIMTRDHTKMLQGFGANLTVETDADGVRTIRLEGRGLTGOVID 240
 |||||
 Db 253 llaglntpgitvtviepimtrdhtekmlqfganltvetdadvrtirlegrgltgovid 312
 |||||

QY 241 VPGDPSSTAFPLVAALLVPGSDVTILNLMPRTTGLILTLQEMGADIEVINPRLAGGED 300
 |||||
 Db 313 vpgdpsstafplvaallvpgsdvtilnlmnprrtgliltlqemgadievinprrlagged 372
 |||||

QY 301 VADLRVRSSTLKGVTVPEDRAPSMIDEPYILAVAAFAEGATVMNGLEELRVKESDRLSA 360
 |||||
 Db 373 vadlrvrstikgvtvpedrapsmideypllavaaaagatvmngleelrvkcsdrlsa 432
 |||||

QY 361 VANGKLINGVDCDEGETSLVVRGRPDGKGLGNASGAATAVTHLDHRIAMSLVMGLVSEN 420
 |||||
 Db 433 vangklngvdcdegetslvvrpgdgkglgnasgaavathldhriamslvmnglvsenp 492
 |||||

QY 421 VTVDDATMIATSPPEFMDLMAGLGAKIELSDTKAA 455
 |||||
 Db 493 vtvdattmiatspefmdlmaglgakielelsdtkaa 527
 |||||

RESULT 4

AAR22300

ID AAR22300 standard; Protein; 455 AA.

XX

AC AAR22300;

XX

DT 03-AUG-1992 (first entry)

XX

DE Class II EPSPS enzyme.

XX

KW Glyphosate tolerant bacteria; herbicides; glyphosate resistant plants;

KW 5'-enolpyruvylshikimate-3-phosphate synthase; weed control.

XX

OS Agrobacterium sp. strain CP4.

XX

PN WO9204449-A.

XX

PD 19-MAR-1992.

XX

PF 28-AUG-1991; 91WO-US06148.

XX

PR 31-AUG-1990; 90US-0576537.

XX

PA (MONS) MONSANTO CO.

XX

PI Barry GF, Kishore GM, Padgett SR;

XX

WPI; 1992-114356/14.

DR

DNA encoding class II 5'-enol pyruvyl shikimate-3-phosphate

PT synthase - for producing plants and bacteria tolerant to

PT glyphosate herbicides

XX

PS Disclosure; Fig 3; 148pp; English.

XX

The sequence is that of the Class II 5'-enolpyruvylshikimate-3

CC phosphate synthase enzyme (EPSPS) of Agrobacterium sp. strain CP4

CC It is used to create glyphosate resistant plants or seeds which

CC can be planted in a field of crops to selectively control weeds.

CC The crops selected for are e.g. corn, wheat, rice, oilseed rape,

CC tobacco and alfalfa. This provides a cost effective, environmentally

CC compatible weed control device. See also AAR22301 and AAR22302.

XX

SQ Sequence 455 AA;

Query Match 99.7%; Score 2282; DB 13; Length 455;
 Best Local Similarity 99.8%; Pred. No. 1.1e-180;
 Matches 454; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLHGASSRPATARKSSGLSGTVPRIKSGTSHRSFPGGLASGETRITGLLEGEDVINTG 60
 |||||
 Db 1 mshgassrpataarkssglsgtvpripksgtshrsfmgfpglasgetritgllegedvintg 60
 |||||

QY 61 KAMQANGARIRKEGDTWIIDGVNGGLLAPEAPLDFGNAATGCRLLTMGLVGVYDFDSTFI 120
 |||||
 Db 61 kamqangarirkegdtwiidgvnggllapeapldfgnaatgcrlltmglvgvydfdstfi 120
 |||||

QY 121 GDASLTKRPMGRVNLPLREMGVQVKSEDDGRLPVTLRGPKTPPIYRVPMSAQVKSAY 180
 |||||

Db	121	gdsaltrkpmgrvlnplrengvgvksedgdrlpvtlrgpktpitryvpmasaqksav	180
Qy	181	LLAGLNTPGITVTETPTWTRDHTKMLQGFGANLTVDADGVRTIRLEGRKLTGQVID	240
Db	181	llaglnpvgitvtetptwtrdhtekmlqgfganltvetdadvrtirlegrkltgqvld	240
Qy	241	VPGDPSTAFPLVAALLVPGSDVTILNVLNMPRTGILILQEMGADIEVINPRLAGGED	300
Db	241	vpgdpsstafplvaallvpgsdvtilnvlmptrtgililtqemgadievlnprlagged	300
Qy	301	VADLRVRSSTLKGVTVPEDRAPSMIDEPILAVAAFAEGATVMNGLEELRVKESDRLSA	360
Db	301	vadlrvrssstlkgvtpedrapsmidvpiilavaaafaegatvmngleelrvkesdrlsa	360
Qy	361	VANGKLNGVDCDEGETSLVVRGPDGKGLGNASGAATAVATHLDRHRIAMSFVLMGLVSENP	420
Db	361	vangklngvdcdegetslvvrpgdgkglgnasgaavathldhriamsflvmglvsenp	420
Qy	421	VTVDATMIATSPFPEFMDLMAGLAKIELSDTKAA	455
Db	421	vtvdatmiatsfpefmdlmaglakieisdtkaa	455
RESULT 5			
ID	AAW34683	standard; Protein; 455 AA.	
AC	AAW34683;		
DT	17-FEB-1998	(first entry)	
DE	Class II EPSP synthase (EPSPS) from Agrobacterium sp. strain CP4.		
KW	5-enolpyruvylshikimate-3-phosphate synthase; EPSPS; glyphosate;		
KW	shikimic acid pathway; phosphoenolpyruvic acid; 3-phosphoshikimic acid;		
KW	5-enolpyruvyl-3-phosphoshikimic acid; herbicide; transgenic plant;		
KW	glyphosate resistance gene; glyphosate-tolerance; promoter.		
OS	Agrobacterium sp. strain CP4.		
FX	Key	Location/Qualifiers	
FT	Region	200..204	
FT	/label=	characteristic_region	
FT	/note=	"see AAW34690"	
FT	Region	26..29	
FT	/label=	characteristic_region	
FT	/note=	"see AAW34691"	
FT	Region	173..177	
FT	/label=	characteristic_region	
FT	/note=	"see AAW34692"	
FT	Region	271..274	
FT	/label=	characteristic_region	
FT	/note=	"see AAW34693"	
PN	US5627061-A.		
XX			
PD	06-MAY-1997.		
XX			
PF	07-JUN-1995;	95US-0476008.	
XX			
PR	13-SEP-1994;	94US-0306063.	
PR	31-AUG-1990;	90US-0576537.	
PR	28-AUG-1991;	91US-0749611.	
XX			
PA	(MONS) MONSANTO CO.		
XX			
PI	Barry GF, Kishore GM, Padgett SR, Stallings WC;		
XX			
DR	WPI; 1997-271315/24.		
DR	N-PSDB; AAT93788.		
XX			
PT	Production of glyphosate-herbicide tolerant plants - using DNA		

PT	encoding class II 5-enol:pyruvyl:shikimate-3-phosphate synthase				
PT	enzyme(s)				
XX	Claim 7; Fig 3; 15lpp; English.				
XX	AAW34683-89 represent a new class of glyphosate-tolerant				
XX	5-enolpyruvylshikimate-3-phosphate synthases (EPSPS). These novel				
CC	EPSPS enzymes have little homology with known Class I EPSPS enzymes, and				
CC	belong to a new class, Class II. The present sequence is from				
CC	Agrobacterium sp. strain CP4. The EPSPS enzymes are part of the shikimic				
CC	acid pathway, which leads to the biosynthesis of aromatic compounds.				
CC	EPSPS converts phosphoenolpyruvic acid (PEP) and is inhibited by the				
CC	to 5-enolpyruvyl-3-phosphoshikimic acid, and is inhibited by the				
CC	herbicide glyphosate. It would be useful to produce transgenic crops				
CC	containing glyphosate resistance genes so that glyphosate-containing				
CC	herbicides can be applied to selectively kill weeds. The novel EPSPS				
CC	enzymes exhibit a low Km for PEP and a high Ki for glyphosate, such that				
CC	when introduced into a plant, the plant is made glyphosate-tolerant, and				
CC	EPSPS enzyme activity is not affected. These class II EPSPS enzymes are				
CC	fused to a chloroplast transit peptide to target the protein into the				
CC	chloroplast, which is the site for the shikimic acid pathway. In				
CC	addition, the EPSPS gene is cloned into a plant under the control of a				
CC	promoter such as figwort mosaic virus promoter or the cauliflower mosaic				
CC	virus promoter, so that expression is enhanced.				
XX					
SQ	Sequence	455 AA;			
Query Match 99.7%; Score 2282; DB 18; Length 455;					
Best Local Similarity 99.8%; Pred. No. 1.1e-180;					
Matches 454; Conservative 0; Mismatches 1; Indels 0; Gaps 0;					
QY	1	MLHGASSRPATARKSGSLSGTVIRIPGDKSIHSRSMFGGLASGETRITGLLEGEDVINTG	60		
DB	1	mshgassrpataarksgslsgtviripgdksihsrsmfmgglasgetritglllegedvintg	60		
QY	61	KAMQAMGARIRKEGDTWIIDGVNGGLLAPAPLDFGNAATGCRLTGVLGVYDFDSTFI	120		
DB	61	kamqamgarirkegdtwiidgvnggllapeapldfngaagcrltmglvgyvdfdstfi	120		
QY	121	GDASLTRKPMGRVLNPLREMGVQVKSEGDRLPVTLRGPKTPTITRYVPNASQAQKSAV	180		
DB	121	gdasltrkpmgrvlnplremgvqvksedgdrlpvtlrgpktptityrvpmasaqvksav	180		
QY	181	LLAGLNTPGITVTETPTWTRDHTKMLQGFGANLTVDADGVRTIRLEGRKLTGQVID	240		
DB	181	llaglntpgittvtetptwtrdhtekmlqgfganltvetdadvrtirlegrkltgqvld	240		
QY	241	VPGDPSTAFPLVAALLVPGSDVTILNVLNMPRTGILILQEMGADIEVINPRLAGGED	300		
DB	241	vpgdpsstafplvaallvpgsdvtilnvlmptrtgililtqemgadievlnprlagged	300		
QY	301	VADLRVRSSTLKGVTVPEDRAPSMIDEPILAVAAFAEGATVMNGLEELRVKESDRLSA	360		
DB	301	vadlrvrsstlkgvtpedrapsmidvpiilavaaafaegatvmngleelrvkesdrlsa	360		
QY	361	VANGKLNGVDCDEGETSLVVRGPDGKGLGNASGAATAVATHLDRHRIAMSFVLMGLVSENP	420		
DB	361	vangklngvdcdegetslvvrpgdgkglgnasgaavathldhriamsflvmglvsenp	420		
QY	421	VTVDATMIATSPFPEFMDLMAGLAKIELSDTKAA	455		
DB	421	vtvdatmiatsfpefmdlmaglakieisdtkaa	455		
RESULT 6					
ID	AAW24474				
XX	AAW24474 standard; Protein; 455 AA.				
AC	AAW24474;				
XX					
DT	02-OCT-1997 (first entry)				
XX					

genomic DNA encoding 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) enzyme referred to as class II EPSPS enzyme which is tolerant to glyphosate. EPSPS genes are useful in producing transgenic bacteria and transgenic plants which are tolerant to glyphosate herbicide. The probe is useful for identifying the presence of a target genomic DNA encoding a EPSPS enzyme. The present sequence is *Agrobacterium* sp. strain CP4 class II EPSPS protein.

Note: The present sequence, SED ID NO: 3 is stated as amino acid sequence throughout the specification. However, it is referred as probe in claim 3 of the specification.

Query Match 99.7%; Score 2282; DB 22; Length 455;
Best Local Similarity 99.8%; Pred. NO. 1.1e-180;
Matches 454; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

CC 1 mshgassrpataarkssglsgrtvpipgdkshrsfmfggiasgetritgllegedvintg 60
CC 61 KAMQAMGARIRKEGDTWIIDGVNGGLLAPAPLDGNAATGCRLLTMGLVGVYDFDSTFI 120
CC 61 kamqamgarirkegdtwiiidvgnggllapeapldfnaatgcrlltmglvgvydfdstfi 120
CC 121 GDASLTRKPMGRVNLPIREMVGQVKSEGDRLPVTLRGPKTPPTITRVPMASQVKSAV 180
CC 121 gdasltrkprmgvrnlplremvgqvksedgdrplvtlrgpktpptityrvpmasqvksav 180
CC 181 LLAGLNTPGITTVIEPIIMTRDHTKMLQGGFANLTVDADGVRTIRLEGRGKLTGQVID 240
CC 181 llaglintpgittviepiimtrdhtekmlqggfnganltvetdadvrtirlegrgkltgqv 240
CC 241 VPGDPSSTAFPLVAALLVPGSDVTILNLMNPTRTGLTLTLOEMGADIEVNPRLAGGED 300
CC 241 vpgdpsstafplvaallvpgsdvtilnlmnpstrtgliltlqemgadievlnprlagged 300
CC 301 VADLRVSRSTLKGVTVPEDRAPSMIDEPILAVAAFAEGATVMNGLEELRVKESDRLSA 360
CC 301 vadlrvrstlkgvtvpedrapsmidelpilavaafaegatvmngleelrvkesdrlsa 360
CC 361 VANGKLKNGVDCDEGETSLVVRGRPDGKGLGNASGAAYATHLDHRIAMSFVLMGLVSEN 420
CC 361 vangklkngvdcdegetslvvrgrpdgkglgnasgaayathldhriamsflvmglvsen 420
CC 421 VTVDATMIATSFPEFMDLMAGLAKIELSDTKAA 455
CC 421 vtvdattmiatsfpefmdlmaglakielsdtkaa 455

CC genomic DNA encoding 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) enzyme referred to as class II EPSPS enzyme which is tolerant to glyphosate. EPSPS genes are useful in producing transgenic bacteria and transgenic plants which are tolerant to glyphosate herbicide. The probe is useful for identifying the presence of a target genomic DNA encoding a EPSPS enzyme. The present sequence is *Agrobacterium* sp. strain CP4 class II EPSPS protein.

Note: The present sequence, SED ID NO: 3 is stated as amino acid sequence throughout the specification. However, it is referred as probe in claim 3 of the specification.

Query Match 99.7%; Score 2282; DB 22; Length 455;
Best Local Similarity 99.8%; Pred. NO. 1.1e-180;
Matches 454; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
DB 1 mshgassrpataarkssglsgrtvpipgdkshrsfmfggiasgetritgllegedvintg 60
QY 61 KAMQAMGARIRKEGDTWIIDGVNGGLLAPAPLDGNAATGCRLLTMGLVGVYDFDSTFI 120
DB 61 kamqamgarirkegdtwiiidvgnggllapeapldfnaatgcrlltmglvgvydfdstfi 120
QY 121 GDASLTRKPMGRVNLPIREMVGQVKSEGDRLPVTLRGPKTPPTITRVPMASQVKSAV 180
DB 121 gdasltrkprmgvrnlplremvgqvksedgdrplvtlrgpktpptityrvpmasqvksav 180
QY 181 LLAGLNTPGITTVIEPIIMTRDHTKMLQGGFANLTVDADGVRTIRLEGRGKLTGQVID 240
DB 181 llaglintpgittviepiimtrdhtekmlqggfnganltvetdadvrtirlegrgkltgqv 240
QY 241 VPGDPSSTAFPLVAALLVPGSDVTILNLMNPTRTGLTLTLOEMGADIEVNPRLAGGED 300
DB 241 vpgdpsstafplvaallvpgsdvtilnlmnpstrtgliltlqemgadievlnprlagged 300
QY 301 VADLRVSRSTLKGVTVPEDRAPSMIDEPILAVAAFAEGATVMNGLEELRVKESDRLSA 360
DB 301 vadlrvrstlkgvtvpedrapsmidelpilavaafaegatvmngleelrvkesdrlsa 360
QY 361 VANGKLKNGVDCDEGETSLVVRGRPDGKGLGNASGAAYATHLDHRIAMSFVLMGLVSEN 420
DB 361 vangklkngvdcdegetslvvrgrpdgkglgnasgaayathldhriamsflvmglvsen 420
QY 421 VTVDATMIATSFPEFMDLMAGLAKIELSDTKAA 455
DB 421 vtvdattmiatsfpefmdlmaglakielsdtkaa 455

RESULT 9
AAE05053 standard; Protein; 455 AA.
AC AAE05053;
DT 10-SEP-2001 (first entry)
XX *Agrobacterium* sp. strain CP4 class II EPSPS protein.
DE *Agrobacterium* sp. strain CP4 class II EPSPS protein.
KW 5-enolpyruvylshikimate-3-phosphate synthase; EPSPS; glyphosate; transformed bacteria; transgenic plant; herbicide.
XX *Agrobacterium* sp. CP4.
OS *Agrobacterium* sp. CP4.
XX US6248876-B1.
PN 19-JUN-2001.
XX 20-AUG-1998; 98US-0137440.
XX 13-SEP-1994; 94US-0306063.
PR 07-APR-1997; 97US-0833485.
PR 31-AUG-1990; 90US-0576537.
PR 28-AUG-1991; 91US-0749611.
XX (MONS) MONSANTO CO.
XX Barry GF, Kishore GM, Padgett SR, Stallings WC;
XX WPI: 2001-407326/43.
DR N-PSDB; AAD0975A.
XX DNA probe capable of use in a polymerase chain reaction for identifying the presence of a target genomic DNA encoding a
PT 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) enzyme -
XX Claim 3; Fig 3; 152pp; English.
XX The present invention relates to a DNA probe capable of use in a
CC polymerase chain reaction for identifying the presence of a target

```
XX PA (MONS ) MONSANTO CO.
XX PI Barry GF, Kishore GM, Padgett SR;
XX PN WPI; 1992-114356/14.
XX PD DNA encoding class II 5'-enol pyruvyl shikimate-3-phosphate
XX PF synthase - for producing plants and bacteria tolerant to
XX PR glyphosate herbicides
XX PS Disclosure; Fig 5; 148pp; English.
XX PT The sequence is that of the Class II 5'-enolpyruvylshikimate-3
XX PH phosphate synthase enzyme (EPSPS) of Achromobacter sp. strain LBAA
XX CC It is used to create glyphosate resistant plants or seeds which
XX CC can be planted in a field of crops to selectively control weeds.
XX CC The crops selected for are e.g. corn, wheat, rice, oilseed rape,
XX CC tobacco and alfalfa. This provides a cost effective, environmentally
XX CC compatible weed control device. See also AAR22300 and AAR22302.
XX SQ Sequence 449 AA;

Query Match 83.1%; Score 1900.5; DB 13; Length 449;
Best Local Similarity 82.9%; Pred. No. 4.3e-149;
Matches 373; Conservative 32; Mismatches 42; Indels 3; Gaps 1;

QY 1 MLHGASRRPATAKSSGLSGTVIRPGDKSISHRSFMPGGLASGETRITGLLEGEDVINTG 60
Db 1 mshsaqpkpatarrsealtgeiripgdkshrsfmgfglasgetritgllgedvintg 60

QY 61 KAMQAMGARIRKEGDWIIDGVNGGLLAPLDFGNATGRLTMGLVGVYDFDSTFI 120
Db 61 ramqamgarkirkegdwiiingvngcllqpeaaldfgnagtgarlmglvgtymktsfi 120

QY 121 GDASTLTKRPMGRVNLPLREMGVQVSKEDGRLPVTLRGPKTPPTITRVPMAQAQVKS 180
Db 121 gdaslkrpnmgrvnlplremgvqveaadgrmpitligkptanpityrvpmsaqvksav 180

QY 181 LLAGLNTPGITVTIEPTMRDHTKMLQFGANLTVETDADGVRTIRLEGRGLTGOVID 240
Db 181 llaglntpgittvtieptmrdrhtkmlqfganltvetdadtgdgvrhritgqgklvgqtlid 240

QY 241 VPGDPSSTAPPLVAALLVPGSDVITILNVLNPNRTTGLILTLQEMGADIEVINPRLAGGED 300
Db 241 vpgdpsstapplvaallvpgsdvitrnlvlnpnrttgliltlqemgadielvinarlagged 300

QY 301 VADLRVRSSTLKGVTYPEDRAPSMIDEPYILAVAAAFAGCATVMNGLEELRVKESDRLSA 360
Db 301 vadlrvrasklkgvtypperapsmideypylaiaaafagcatvmdgldelrvkesdrlaa 360

QY 361 VANGLKLVGDCDEGETSLVWRGPRPGKGLGNASGAATAVTHLDRHTAMSPFLVGLVSENP 420
Db 361 vangleangvdcdegetslvwrgrppgkglgnasgaavathldhrtamsflvnglvsenp 420

QY 421 VTVDATMIATSPPEFMDLMAGLAKIELS 450
Db 418 vtvdnsmiatsfpefmdmmpglgakiels 447

RESULT 10
AAR22302
ID AAR22302 standard; Protein; 449 AA.
XX AC AAR22302;
XX DT 03-AUG-1992 (first entry)
XX DE Class II EPSPS enzyme.
XX KW glyphosate tolerant bacteria; herbicides; glyphosate resistant plants;
KW 5'-enolpyruvylshikimate-3-phosphate synthase; weed control.
```

```
XX OS Pseudomonas sp. strain PG2982.
XX PN WO9204449-A.
XX PD 19-MAR-1992.
XX PF 28-AUG-1991; 91WO-US06148.
XX PR 31-AUG-1990; 90US-0576537.
XX PA (MONS ) MONSANTO CO.
XX PI Barry GF, Kishore GM, Padgett SR;
XX PN WPI; 1992-114356/14.
XX PD DNA encoding class II 5'-enol pyruvyl shikimate-3-phosphate
XX PF synthase - for producing plants and bacteria tolerant to
XX PR glyphosate herbicides
XX PS Disclosure; Fig 7; 148pp; English.
XX PT The sequence is that of the Class II 5'-enolpyruvylshikimate-3
XX PH phosphate synthase enzyme (EPSPS) of Pseudomonas sp. strain PG2982
XX CC It is used to create glyphosate resistant plants or seeds which
XX CC can be planted in a field of crops to selectively control weeds.
XX CC The crops selected for are e.g. corn, wheat, rice, oilseed rape,
XX CC tobacco and alfalfa. This provides a cost effective, environmentally
XX CC compatible weed control device. See also AAR22300 and AAR22301.
XX SQ Sequence 449 AA;

Query Match 83.1%; Score 1900.5; DB 13; Length 449;
Best Local Similarity 82.9%; Pred. No. 4.3e-149;
Matches 373; Conservative 32; Mismatches 42; Indels 3; Gaps 1;

QY 1 MLHGASRRPATAKSSGLSGTVIRPGDKSISHRSFMPGGLASGETRITGLLEGEDVINTG 60
Db 1 mshsaqpkpatarrsealtgeiripgdkshrsfmgfglasgetritgllgedvintg 60

QY 61 KAMQAMGARIRKEGDWIIDGVNGGLLAPLDFGNATGRLTMGLVGVYDFDSTFI 120
Db 61 ramqamgarkirkegdwiiingvngcllqpeaaldfgnagtgarlmglvgtymktsfi 120

QY 121 GDASTLTKRPMGRVNLPLREMGVQVSKEDGRLPVTLRGPKTPPTITRVPMAQAQVKS 180
Db 121 gdaslkrpnmgrvnlplremgvqveaadgrmpitligkptanpityrvpmsaqvksav 180

QY 181 LLAGLNTPGITVTIEPTMRDHTKMLQFGANLTVETDADGVRTIRLEGRGLTGOVID 240
Db 181 llaglntpgittvtieptmrdrhtkmlqfganltvetdadtgdgvrhritgqgklvgqtlid 240

QY 241 VPGDPSSTAPPLVAALLVPGSDVITILNVLNPNRTTGLILTLQEMGADIEVINPRLAGGED 300
Db 241 vpgdpsstapplvaallvpgsdvitrnlvlnpnrttgliltlqemgadielvinarlagged 300

QY 301 VADLRVRSSTLKGVTYPEDRAPSMIDEPYILAVAAAFAGCATVMNGLEELRVKESDRLSA 360
Db 301 vadlrvrasklkgvtypperapsmideypylaiaaafagcatvmdgldelrvkesdrlaa 360

QY 361 VANGLKLVGDCDEGETSLVWRGPRPGKGLGNASGAATAVTHLDRHTAMSPFLVGLVSENP 420
Db 361 vangleangvdcdegetslvwrgrppgkglgnasgaavathldhrtamsflvnglvsenp 420

QY 421 VTVDATMIATSPPEFMDLMAGLAKIELS 450
Db 418 vtvdnsmiatsfpefmdmmpglgakiels 447

RESULT 11
AAR22302
```

AAW34684 standard; Protein; 449 AA.

AAW34684;

17-FEB-1998 (first entry)

Class II EPSP synthase (EPSPS) from *Achromobacter* sp. strain LBAA.

5-enolpyruvylshikimate-3-phosphate synthase; EPSPS; glyphosate; shikimic acid pathway; phosphoenolpyruvic acid; 3-phosphoshikimic acid; 5-enolpyruvyl-3-phosphoshikimic acid; herbicide; transgenic plant; glyphosate resistance gene; glyphosate-tolerance; promoter.

Achromobacter sp. strain LBAA.

Key Location/Qualifiers

Region 200..204

/label= characteristic_region

/note= "see AAW34690"

Region 26..29

/label= characteristic_region

/note= "see AAW34691"

Region 173..177

/label= characteristic_region

/note= "see AAW34692"

Region 271..274

/label= characteristic_region

/note= "see AAW34693"

US5627061-A.

06-MAY-1997. 95US-0476008.

13-SEP-1994; 94US-0306063.

31-AUG-1990; 90US-0576537.

28-AUG-1991; 91US-0749611.

(MONS) MONSANTO CO.

Barry GF, Kishore GM, Padgett SR, Stallings WC;

WPI; 1997-271315/24.

N-PSDB; AAT93789.

Production of glyphosate-herbicide tolerant plants - using DNA encoding class II 5-enolpyruvylshikimate-3-phosphate synthase enzyme(s)

Disclosure; Fig 4; 151pp; English.

AAW34683-89 represent a new class of glyphosate-tolerant 5-enolpyruvylshikimate-3-phosphate synthases (EPSPS). These novel EPSPS enzymes have little homology with known Class I EPSPS enzymes, and belong to a new class, Class II. The present sequence was isolated from *Achromobacter* sp. strain LBAA. The EPSPS enzymes are part of the shikimic acid pathway, which leads to the biosynthesis of aromatic compounds. EPSPS converts phosphoenolpyruvic acid (PEP) and 3-phosphoshikimic acid to 5-enolpyruvyl-3-phosphoshikimic acid, and is inhibited by the herbicide glyphosate. It would be useful to produce transgenic crops containing glyphosate resistance genes so that glyphosate-containing herbicides can be applied to selectively kill weeds. The novel EPSPS enzymes exhibit a low Km for PEP and a high Ki for glyphosate, such that when introduced into a plant, the plant is made glyphosate-tolerant, and EPSPS enzyme activity is not affected. These class II EPSPS enzymes are fused to a chloroplast transit peptide to target the protein into the chloroplast, which is the site for the shikimic acid pathway. In addition, the EPSPS gene is cloned into a plant under the control of a promoter such as figwort mosaic virus promoter or the cauliflower mosaic virus promoter, so that expression is enhanced.


```
Query Match      83.1%; Score 1900.5; DB 18; Length 449;
Best Local Similarity 82.9%; Pred. No. 4.3e-149;
Matches 373; Conservative 32; Mismatches 42; Indels 3; Gaps 1;

QY 1 MLHGASSRPATARKSGSLGSLGVTRIPGDKSISHRSPFPGGLASGETRITGLLEGEDVINTG 60
Db 1 mshsaaspkpatarrsealtgeiripgdksishrsfmfgglasgetritgllgedvintg 60
QY 61 KAMQAMCARIKKEDTWIIDGVNGGLLAPPEALDFGNAATGCLTMGLGVVDVDFSTFI 120
Db 61 ramqamgakarkegdvwlvgngngcillqpeaaldfgnagtgarltnglvgtymktsfi 120
QY 121 GDASLTRKPRGRVLNPLREMGVQVKSEGDRLPVLTRGPKTPTPTITVYRVPMAASQVKS AV 180
Db 121 gdaalskrpmgrvlnplremngvqveaadgrmpitllgpktpantpityrvpmasqvksav 180
QY 181 LLAGLNPFGTITVTEPIWTRDHTKMLQGGCANLTVETDAGVGTIRLEGRGKLTGOVID 240
Db 181 llaglnpfgvtvtepiwtrdhtekmlqggfagadltvetdkdgvzhrfritggkllvggtid 240
QY 241 VPGDPSSTAFPLVAALLVPGSDVTILNVLNMPTRTGLILTLQEMGADIEVINPRLAGGED 300
Db 241 vpgdpsstaflvaallvegsvdvtirnlvlnmptrtglltllqemgadievlnarlagged 300
QY 301 VADLRVRSSTLKGVTVPEDRAPSMIDEPYILAAFAAEGATVMNGLEELRVKESDRLSA 360
Db 301 vadlrvrasklkgvvvppeapsmideypvllaasfaegetvmdgidelrvkesdrila 360
QY 361 VANGKLNGVDCDEGETSLVVRGRPDGKGLGNASGAATAVATHDHRAMSLFVLVGLVSEN 420
Db 361 vangleangvdcdegetslvvrgrpdgkglgnasgaatavathdhriamsflvmglvae 420
QY 421 VTDDATMIATSFPEFMDLMAGLGAKELS 450
Db 418 vtvdssnmatsfpefmdmpgigakieLS 447

RESULT 15
AAW71611
ID AAW71611 standard; Protein; 449 AA.
XX
AC AAW71611;
XX
DT 14-DEC-1998 (first entry)
XX
DE Pseudomonas sp. strain PG2982 Class II EPSPS.
XX
KW 5-enolpyruvylshikimate-3-phosphate synthase; class II EPSPS enzyme;
KW glyphosate herbicide; transformed bacteria; class I EPSPS enzyme;
KW resistance; inhibition; 5-enolpyruvyl-3-phosphoshikimic acid;
KW N-phosphonomethylglycine.
XX
OS Pseudomonas sp.
XX
PN US5804425-A.
XX
PD 08-SEP-1998.
XX
PF 07-APR-1997; 97US-0833485.
XX
PR 13-SEP-1994; 94US-0306063.
PR 31-AUG-1990; 90US-0576537.
PR 28-AUG-1991; 91US-0749611.
PR 07-APR-1997; 97US-0833485.
XX
PA (MONS ) MONSANTO CO.
XX
PI Barry GF, Kishore GM, Padgett SR, Stallings WC;
XX
DR WPI; 1998-505657/43.
XX
DR N-PSDB; AAV58012.
XX
PT Glyphosate resistant 5-enolpyruvylshikimate-3-phosphate synthase -
```

```
QY 181 LLAGLNPFGTITVTEPIWTRDHTKMLQGGCANLTVETDAGVGTIRLEGRGKLTGOVID 240
Db 181 llaglnpfgvtvtepiwtrdhtekmlqggfagadltvetdkdgvzhrfritggkllvggtid 240
QY 241 VPGDPSSTAFPLVAALLVPGSDVTILNVLNMPTRTGLILTLQEMGADIEVINPRLAGGED 300
Db 241 vpgdpsstaflvaallvegsvdvtirnlvlnmptrtglltllqemgadievlnarlagged 300
QY 301 VADLRVRSSTLKGVTVPEDRAPSMIDEPYILAAFAAEGATVMNGLEELRVKESDRLSA 360
Db 301 vadlrvrasklkgvvvppeapsmideypvllaasfaegetvmdgidelrvkesdrila 360
QY 361 VANGKLNGVDCDEGETSLVVRGRPDGKGLGNASGAATAVATHDHRAMSLFVLVGLVSEN 420
Db 361 vangleangvdcdegetslvvrgrpdgkglgnasgaatavathdhriamsflvmglvae 420
QY 421 VTDDATMIATSFPEFMDLMAGLGAKELS 450
Db 418 vtvdssnmatsfpefmdmpgigakieLS 447

RESULT 14
AAW24480
ID AAW24480 standard; Protein; 449 AA.
XX
AC AAW24480;
XX
DT 02-OCT-1997 (first entry)
XX
DE Class II EPSPS used for glyphosate resistant plant production.
XX
KW 5-enolpyruvylshikimate-3-phosphate synthase; EPSPS; Class II;
KW glyphosate resistant; transgenic plant; herbicide; shikimic acid;
KW fusion protein; soybean; corn; wheat; barley; cotton; sugarbeet;
KW tobacco.
XX
OS Pseudomonas sp. strain PG2982.
XX
PN US5633435-A.
XX
PD 27-MAY-1997.
XX
PF 31-AUG-1990; 90US-0576537.
XX
PR 13-SEP-1994; 94US-0306063.
PR 31-AUG-1990; 90US-0576537.
PR 28-AUG-1991; 91US-0749611.
XX
PA (MONS ) MONSANTO CO.
XX
PI Barry GF, Kishore GM, Padgett SR, Stallings WC;
XX
DR WPI; 1997-297418/27.
XX
DR N-PSDB; AAW77316.
XX
PT New isolated 5-enolpyruvylshikimate-3-phosphate synthase gene -
PT used for transforming plants to produce plants which are tolerant to
PT glyphosate herbicide
XX
PS Claim 7; Column 69-72; 154pp; English.
XX
CC AAW24480 shows the sequence of a class II 5-enolpyruvylshikimate-3-
CC phosphate synthase (EPSPS) enzyme. Class II EPSPS enzymes are tolerant
CC to glyphosate herbicides. EPSPS and sequences encoding it are used for
CC the production of herbicide resistant (glyphosate-tolerant) plants
CC such as corn, wheat, rice, barley, soybean, cotton, sugarbeet, oilseed
CC rape, canola, flax, sunflower, potato, tobacco, alfalfa, poplar,
CC pine, eucalyptus, apple, lettuce, peas, lentils, grape and turf grasses.
XX
SQ Sequence 449 AA;
```


PT useful for characterisation of the enzyme to determine inhibition
PT data values
XX
PS Disclosure; Fig 5; 152pp; English.
XX
XX An isolated 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) has the
CC sequence domains (I), (II), (III) and (IV): -R-X1-H-X2-E- (I),
CC -G-D-K-X3- (II); -S-A-O-X4-K- (III); and -N-X5-T-R- (IV). Where X1 = G,
CC S, T, C, Y, N, Q, D or E; X2 and X3 = S or T; X4 and X5 = A, R, N, D, C,
CC Q, E, G, H, I, L, K, M, F, P, S, T, W, Y or V. The EPSPS enzyme,
CC produced by recombinant methods, can be used in kinetic studies to
CC determine Ki and Km values of the enzyme for its characterisation. The
CC enzyme is normally used for the production of 5-enolpyruvyl-3-
CC phosphoshikimic acid in plants, and most forms of the enzyme are
CC inhibited by N-phosphonomethylglycine ('glyphosate') herbicides.
CC Inhibition data enables more accurate values of concentrations of
CC herbicide to be used when growing the plant without being detrimental to
CC it. This enables the plant to be grown in the presence of the herbicide,
CC being used to inhibit the growth of undesired plants. The present
CC sequence represents a Class II EPSPS from bacterial isolate
CC Pseudomonas sp. strain PG2982.
XX
SQ Sequence 449 AA;

Query Match 83.1%; Score 1900.5; DB 19; Length 449;
Best Local Similarity 82.9%; Pred. No. 4.3e-149;
Matches 373; Conservative 32; Mismatches 42; Indels 3; Gaps 1;

Qy 1 MLHGASSRPATARKSSGLSGTVIRPGDKSISHSRPFMGGLASGETRITGLLEGEDVINTG 60
Dy 1 mshasaspkatarrealtgeiripgdkshsrfmgfglasgetritgllgegedvintg 60

Qy 61 KAMQAMGARIRKEGDTIIDGVNGGGLLAPEAPLDGNAATGCRLTMLGLVGYDFDSTFI 120
Dy 61 ramqamgakirkegdviiingvngcgllqpeaaldfnagtgartmgltvgtydmktsfi 120

Qy 121 GDASLTRPMGRVNLPLREMGVQKSESDGRLPVTLRGPKTPTITYRVPWASAQVKSAY 180
Dy 121 gdaalskrpmgrvnlplremgvqveaadgdrmpiltligpkcanpityrvpmasaqvksav 180

Qy 181 LLAGINTPGITVTIEPIMTRDHTKMLQGFGANLTVDADGVRTIRLEGRKLTGQVID 240
Dy 181 llagintpgvttviepvmtrdhtekmlqgfgadltvetdkdgvhrhritggklyvgqtid 240

Qy 241 VPGDPSSTAFPLVALLVPGSDVTILNVMNPTRTGLTLTQEMGADIEVINPRLAGGED 300
Dy 241 vpgdpsstafplvaallvegsdvtirnlmnptrtgltiltqemgadievlnarlagged 300

Qy 301 VADLRVRSSTLKGVTVPEDRAPSMIDEPYILAVAAFAEGATVMNGLEELRVKESDRLSA 360
Dy 301 vadlrvrasklkgvvppperapsmdiepyilaasfaegetvmdgldelrvkesdrila 360

Qy 361 VANGILKLVGDCDEGETSLVYRGRPDGKGLGNASGAAVATHLDHRIAMSFVLMGLVSENP 420
Dy 361 vangleangvdcetegemslvtrgpdgkglg---ggvathldhriamsflvmglvsekp 420

Qy 421 VTVDATWIAATSPFEFMDLMAGLAKIELS 450
Dy 418 vtvdnsniatsfpfmdmmpglgakiels 447

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 15, 2002, 13:59:21 ; Search time 17.59 Seconds

(without alignments)
1001.557 Million cell updates/sec

Title: US-09-464-099A-3

Perfect score: 2288

Sequence: 1 MSHGASRRPATARKSSGLSG.....FMDLMACLGAKIELSDTKAA 455

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

arched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2288	100.0	455	1 AROA_AGRSP	Q9r4e4 agrobacteri
2	1906.5	83.3	449	1 AROA_PSES2	P56952 pseudomonas
3	884	38.6	431	1 AROA_BACHD	Q9kca6 bacillus ha
4	867.5	37.9	447	1 AROA_SYNV3	Q59975 synecocyst
5	806.5	35.2	430	1 AROA_LACLA	Q9ceuo lactococcus
6	806	35.2	443	1 AROA_BACNO	Q46550 bacteroides
7	803.5	35.1	427	1 AROA_STRPN	Q9s400 streptococc
8	802.5	35.1	430	1 AROA_LACLC	P20691 bacillus su
9	785	34.3	428	1 AROA_BACSU	Q67494 aquifex ae
10	766.5	33.5	431	1 AROA_AQUAE	Q9wy10 thermotoga
11	695	30.4	410	1 AROA_THEMEA	P53112 campylobact
12	590	25.8	428	1 AROA_CAMJE	Q03615 staphylococ
13	577	25.2	430	1 AROA_STAAD	Q92k17 helicobacte
14	521.5	22.8	429	1 AROA_HELPJ	P56197 helicobacte
15	518.5	22.7	429	1 AROA_HELPJ	O26860 methanobact
16	496	21.7	419	1 AROA_METTH	Q57925 methanococc
17	459.5	20.1	429	1 AROA_METTA	Q9rvd3 deinococcus
18	408	17.8	439	1 AROA_DEIRA	O28775 archaeoglob
19	403	17.6	416	1 AROA_ARCFU	Q9yek9 aeropyrum p
20	402.5	17.6	427	1 AROA_AERPE	P19688 versinia en
21	402.5	17.6	427	1 AROA_YEREN	O60112 versinia pe
22	395	17.3	428	1 AROA_YERPE	P53310 haemophilus
23	387.5	16.9	432	1 AROA_HAESO	Q07638 escherichia
24	378	16.5	427	1 AROA_ECOLI	Q9zff7 shigella so
25	376	16.4	427	1 AROA_SHISO	Q9v1h1 pyrococcus
26	373.5	16.3	410	1 AROA_PYRAB	P54220 pasteurella
27	373	16.3	432	1 AROA_PASHA	P24497 klebsiella
28	372	16.3	427	1 AROA_KLEPN	Q04570 pasteurella
29	372	16.3	440	1 AROA_PASMO	O87006 shigella dy
30	371	16.2	427	1 AROA_SHIDI	P39915 burkholderi
31	365.5	16.0	463	1 AROA_BURPS	Q9x4h2 edwardsiell
32	365	16.0	428	1 AROA_EDWIC	Q03421 haemophilus
33	363.5	15.9	432	1 AROA_HAEIN	

34	357	15.6	427	1 AROA_SALTY	P07637 salmonella
35	348	15.2	427	1 AROA_SALGL	P22299 salmonella
36	347	15.2	427	1 AROA_SALTI	P19786 salmonella
37	323.5	14.1	516	1 AROA_PETHY	P11043 petunia hyb
38	321.5	14.1	1588	1 ARO1_YEAST	P08566 s pentafunc
39	317	13.9	441	1 AROA_CHLMU	Q9pk28 chlamydia m
40	315.5	13.8	518	1 ARO1_TOBAC	P23981 nicotiana t
41	314.5	13.7	440	1 AROA_CHLTR	O84371 chlamydia t
42	313	13.7	428	1 AROA_BUCAP	Q59178 buchnera ap
43	310.5	13.6	520	1 AROA_LYCES	P10748 lycopersico
44	308.5	13.5	1573	1 ARO1_SCHPO	Q9p7r0 s pentafunc
45	306.5	13.4	445	1 AROA_CHLPN	Q9z6m0 chlamydia p

ALIGNMENTS

RESULT 1
AROA_AGRSP STANDARD; PRT; 455 AA.
AC Q9R4E4;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GN AROA.
OS Agrobacterium sp. (strain Cp4).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Agrobacterium.
OX NCBI_TaxID=361;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-28; 47-61 AND 321-333.
RA Barry G.F., Kishore G.M., Padgett S.R., Stallings W.C.;
RT "Glyphosate-tolerant 5-enolpyruvylshikimate-3-phosphate synthases.";
RL Patent number US5633435, 27-MAY-1997.
RN [2]
RP SEQUENCE OF 1-15.
RX MEDLINE=96182485; PubMed=8598558;
RA Harrison L.A., Bailey M.R., Naylor M.W., Ream J.E., Hammond B.G.,
RA Nida D.L., Burnette B.L., Nickson T.E., Mitsky T.A., Taylor M.L.,
RA Fuchs R.L., Padgett S.R.;
RT "The expressed protein in glyphosate-tolerant soybean, 5-enolpyruvylshikimate-3-phosphate synthase from Agrobacterium sp. strain Cp4, is rapidly digested in vitro and is not toxic to acutely gavaged mice.";
RL J. Nutr. 126:728-740(1996).
CC -!- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate - phosphate + 5'-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -!- BIOTECHNOLOGY: Introduced by genetic manipulation and expressed in glyphosate-tolerant soybean, canola, cotton and maize by Monsanto. Developed to provide new weed-control options for farmers.
CC Expression of this protein in plants imparts high levels of glyphosate tolerance.
CC -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
DR InterPro: IPR001986; EPSP_synthase.
DR Pfam: PF00275; EPSP_synthase; 1.
DR PROSITE: PD001867; EPSP_synthase; 1.
DR PROSITE: PS00104; EPSP_SYNTHASE_1; 1.
DR PROSITE: PS00885; EPSP_SYNTHASE_2; 1.
KW Aromatic amino acid biosynthesis; Transferrase; Herbicide resistance; Genetically modified food.
FT CONFLICT 2 S -> L (IN REF. 2).
SQ SEQUENCE 455 AA; 47588 MW; 236580D08D6EF422 CRC64;

Query Match 100.0%; Score 2288; DB 1; Length 455;
Best Local Similarity 100.0%; Pred. No. 3.8e-135;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MSHGASSRPATARKSSGLSGTGVIRIPGDKSISHRSFMPFGLASGETRITGLLEGEDVINTG 60
DB 1 MSHGASSRPATARKSSGLSGTGVIRIPGDKSISHRSFMPFGLASGETRITGLLEGEDVINTG 60
QY 61 KAMQAMGARIRKEGDTWIIDGVNGGLLAPLDFGNAATGCLRLTMGLVGVYDFDSTFI 120
DB 61 KAMQAMGARIRKEGDTWIIDGVNGGLLAPLDFGNAATGCLRLTMGLVGVYDFDSTFI 120
QY 121 GDASLTRKPRMGRVNLPLREMGVQVKSSEDGRLPVTLRGPKTPTPTITVRVPMASQAQVSAV 180
DB 121 GDASLTRKPRMGRVNLPLREMGVQVKSSEDGRLPVTLRGPKTPTPTITVRVPMASQAQVSAV 180
QY 181 LLAGLNTPGITTVTIEPIMTRDHTKMLQGFAGNLTVEVDGVRTIRLEGRGKLTGVQVID 240
DB 181 LLAGLNTPGITTVTIEPIMTRDHTKMLQGFAGNLTVEVDGVRTIRLEGRGKLTGVQVID 240
QY 241 VPGDPSSTAFPLVAALLVPGSDVTILNLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
DB 241 VPGDPSSTAFPLVAALLVPGSDVTILNLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
QY 301 VADLRVRSSTLKGVTVPEDRAPSMIDEPYILAVAAAFAGATVNGLEELRVKESDRLSA 360
DB 301 VADLRVRSSTLKGVTVPEDRAPSMIDEPYILAVAAAFAGATVNGLEELRVKESDRLSA 360
QY 361 VANGKLNGVDCDEGETSLVVRGPRDGKGLGNASGAATAVATHLDRHRIAMSFVLMGLVSEN 420
DB 361 VANGKLNGVDCDEGETSLVVRGPRDGKGLGNASGAATAVATHLDRHRIAMSFVLMGLVSEN 420
QY 421 VTVDATMIATSFPEFMDLMAGLAKIELSDTKAA 455
DB 421 VTVDATMIATSFPEFMDLMAGLAKIELSDTKAA 455
```

```
RESULT 2
AROA_PSES2
ID AROA_PSES2 STANDARD; PRT; 449 AA.
AC P56952;
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
DE enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GN AROA.
OS Pseudomonas sp. (strain PG2982), and
OS Achromobacter sp. (strain LBAA).
OC Bacteria; Proteobacteria.
OX NCBI_TaxID=308, 129026;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 2-16.
RA Barry G.F., Kishore G.M., Padgett S.R., Stallings W.C.;
RA "Glyphosate-tolerant 5-enolpyruvylshikimate-3-phosphate synthases.";
RA Patent number US5633435, 27-MAY-1997.
RL
CC -1- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
CC phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC -1- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
CC THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- MISCELLANEOUS: RESISTANT TO THE ANTIBIOTIC GLYPHOSATE.
CC -1- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
DR InterPro: IPR001986; EPSP_Synthase.
DR Pfam: PF00275; EPSP_Synthase; 1.
DR ProDom: PD001867; EPSP_Synthase; 1.
DR PROSITE: PS00104; EPSP_SYNTHASE_1; 1.
DR PROSITE: PS00885; EPSP_SYNTHASE_2; 1.
KW Aromatic amino acid biosynthesis; Transferrase; Herbicide resistance.
SQ SEQUENCE 449 AA; 47297 MW; 447F213EECAEFCL CRC64;
```

Query Match 83.3%; Score 1906.5; DB 1; Length 449;
Best-Local Similarity 83.1%; Pred. No. 1.8e-111;
Matches 374; Conservative 32; Mismatches 41; Indels 3; Gaps 1;

```
QY 1 MSHGASSRPATARKSSGLSGTGVIRIPGDKSISHRSFMPFGLASGETRITGLLEGEDVINTG 60
DB 1 MSHGASSRPATARKSSGLSGTGVIRIPGDKSISHRSFMPFGLASGETRITGLLEGEDVINTG 60
QY 61 KAMQAMGARIRKEGDTWIIDGVNGGLLAPLDFGNAATGCLRLTMGLVGVYDFDSTFI 120
DB 61 KAMQAMGARIRKEGDTWIIDGVNGGLLAPLDFGNAATGCLRLTMGLVGVYDFDSTFI 120
QY 121 GDASLTRKPRMGRVNLPLREMGVQVKSSEDGRLPVTLRGPKTPTPTITVRVPMASQAQVSAV 180
DB 121 GDASLTRKPRMGRVNLPLREMGVQVKSSEDGRLPVTLRGPKTPTPTITVRVPMASQAQVSAV 180
QY 181 LLAGLNTPGITTVTIEPIMTRDHTKMLQGFAGNLTVEVDGVRTIRLEGRGKLTGVQVID 240
DB 181 LLAGLNTPGITTVTIEPIMTRDHTKMLQGFAGNLTVEVDGVRTIRLEGRGKLTGVQVID 240
QY 241 VPGDPSSTAFPLVAALLVPGSDVTILNLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
DB 241 VPGDPSSTAFPLVAALLVPGSDVTILNLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
QY 301 VADLRVRSSTLKGVTVPEDRAPSMIDEPYILAVAAAFAGATVNGLEELRVKESDRLSA 360
DB 301 VADLRVRSSTLKGVTVPEDRAPSMIDEPYILAVAAAFAGATVNGLEELRVKESDRLSA 360
QY 361 VANGKLNGVDCDEGETSLVVRGPRDGKGLGNASGAATAVATHLDRHRIAMSFVLMGLVSEN 420
DB 361 VANGKLNGVDCDEGETSLVVRGPRDGKGLGNASGAATAVATHLDRHRIAMSFVLMGLVSEN 420
QY 421 VTVDATMIATSFPEFMDLMAGLAKIELS 450
DB 421 VTVDATMIATSFPEFMDLMAGLAKIELS 450
QY 418 VTVDATMIATSFPEFMDLMAGLAKIELS 447
DB 418 VTVDATMIATSFPEFMDLMAGLAKIELS 447
RESULT 3
AROA_BACHD
ID AROA_BACHD STANDARD; PRT; 431 AA.
AC O9KCA6;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
DE enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GN AROA OR AROE OR BH1667.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RA "Complete genome sequence of the alkaliphilic bacterium Bacillus
RA halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
CC -1- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
CC phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC -1- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN THE
CC BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC -1- SUBUNIT: Monomer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
```

```

DR EMBL: AP001512; BAB05386.1; -
DR InterPro: IPR00186; ESpP_syntase.
DR Pfam: PF00275; ESpP_syntase; 1.
DR ProDom: PD001867; ESpP_syntase; 1.
DR PROSITE: PS00104; ESpP_SYNTHASE_1; 1.
DR PROSITE: PS00885; ESpP_SYNTHASE_2; 1.
DR Aromatic amino acid biosynthesis; Transferrase; Complete proteome.
KW SEQUENCE 431 AA; 45495 MW; 12F4FBE7BA0743D CRC64;

Query Match 38.6%; Score 884; DB 1; Length 431;
Best Local Similarity 44.2%; Pred. No. 4.7e-48;
Matches 188; Conservative 72; Mismatches 153; Indels 12; Gaps

QY 15 SSGLSGTVRIPGDKSISHRSFMPGGGLASGETRITGLLEGEDVINTKAMQAMCARIKKEG 74
Db 10 AKGLKTKTIVPGDKSISHRAVMFGALAKKTTTVEGFLPGADCLSTSCFQKILGVSTEQAE 69

75 DTWIIDGVNGGGLLAPLEAPLDFGNAAATGRLTMGLVGVYDFDSTFGDASLTKRPMGRVL 134
70 ERYTVKGGWDGLREFSDIILDVNSGTTTWRLLIGLILSTLPFHSVIIGDESIGKRPKRVTT 129

QY 135 NPLREMGVQVKSED - GDRLPVTLRUGPKTPTPTITYRVPMSAQVKSAVLLAGLNTPGITTV 193
Db 130 EPLKSMGAQTDGDRDHGNLTPLSIRGGQL - KGIDFHPVASAQMSAILLAGLRAEGKTSV 188

QY 194 IEPIMTRDHTPEKMLQCGFANLTVETDADGVRTIRLGRCKLTCQVIDVPGDSSSTAFPLV 253
Db 189 TEPAKTNHTMERLMEAFGNVI ---- EKDGL - TVSIEGGQMLTGQHVHVPGDISSAAFFLV 243

QY 254 AALLVPGSDVTILNLVLMNPTRTGLIILTLOEMGADIIVINPRLAGGEDVADLRVRSSTLKG 313
Db 244 AGAMVPHSRITLTVNGINPTFRAGILEVLKQMGATLAMENERVGGEPVADLTLETSVLQG 303

QY 314 VTPEDRAPSNIDEYPIILAAAFAGATVMNGLBELRVKESDRLSAVANGILKNGVDCD 373
Db 304 VEIGGDIIPLRIDEIPILAVLATQASGRVTIKDAESELVKETNRIDTVVYSELTKLGASTH 363

QY 374 EGTSLTVRGRPGDKGLGNASGAAVATHLDHRTAMSLFVNGVLSENPVTVDDATMTATSF 433
Db 364 ATDDGMIIIEGPTPLKG ----- GYTVSSHGDHRIGMANATAALLAEKPVTVTEGTEIAVSY 418

QY 434 PFMD 438
Db 419 PSEFD 423

```

ULT 4
 AROA_SYN3
 ID AROA_SYN3 STANDARD; PR3; 447 AA.
 AC Q59975; Q59974;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE 3-phosphohikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
 DE enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
 DE AROA OR SLR0444.
 GN AROA
 OS Synecocystis sp. (strain PCC 6803).
 OC Bacteria; Cyanobacteria; Chroococcales; Synecocystis.
 OX NCBI_TaxID=1148;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94299161; PubMed=8026753;
 RX dalla Chiesa M., Mayes S.R., Maskell D.S., Nixon P.J., Barber J.;
 RT "An arao homologue from Synecocystis sp. PCC 6803.";
 RL Gene 144:145-146(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97061201; PubMed=8905231;
 RX Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
 RA Miyajima N., Hiroseawa M., Sugitara M., Sasamoto S., Kimura T.,
 RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
 RA Shimpou S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,

```

RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
RN [3]
RX SEQUENCE OF 103-194 FROM N.A.
RT MEDLINE=93307506; PubMed=7686511;
RT Mayes S.R., dalla Chiesa M., Zhang Z., Barber J.;
RT "The genes aroA and trnQ are located upstream of psbO in the
RT chromosome of Synechocystis 6803.";
RT FEBS Lett. 325:253-261(1993).
CC -1- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
CC phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC -1- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
CC THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J75325; CAA53074.1; -
CC EMBL; D90914; BAA18477.1; -
CC DR EMBL; J72784; CAA51291.1; -
CC DR EMBL; J72784; CAA51291.1; -
CC DR InterPro: IPR001986; EPSP_synthase.
CC DR Pfam: PF00275; EPSP_synthase; 1.
CC DR ProDom: PD001867; EPSP_synthase; 1.
CC DR PROSITE; PS00104; EPSP_SYNTHASE.1; 1.
CC DR PROSITE; PS00895; EPSP_SYNTHASE.2; 1.
CC DR Aromatic amino acid biosynthesis; Transferase; Complete proteome.
CC SEQUENCE 447 AA; 47046 MW; 2F3C8AD26B5A7BCE CRC64;
SQ

```

[illegible]

[illegible]

	Db	436 EF 437	
		II	
		RESULT 5	
		ARO_A_LACLA	
		ID AROA_LACLA STANDARD; PRT; 430 AA.	
		AC O9CEUO;	
		DT 16-OCT-2001 (Rel. 40, Created)	
		DT 16-OCT-2001 (Rel. 40, Last sequence update)	
		DT 01-MAR-2002 (Rel. 41, Last annotation update)	
		DE 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-	
		enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).	
		ARO OR LL1744.	
		GN Lactococcus lactis (subsp. lactis) (Streptococcus lactis).	
		OS Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;	
		Lactococcus.	
		OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;	
		OX NCBI_TaxID=1360;	
		[1]	
		SEQUENCE FROM N.A.	
		RN STRAIN=ILI1403;	
		RP MEDLINE=21235186; PubMed=11337471;	
		RC Bolotin A., Wincker P., Mauger S., Jalllon O., Malarne K.,	
		Weissenbach J., Ehrlich S.D., Sorokin A.;	
		RA "The complete genome sequence of the lactic acid bacterium Lactococcus	
		lactis ssp. lactis IL1403."	
		RT Genome Res. 11:731-753(2001).	
		RL -1- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =	
		phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.	
		CC -1- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN	
		THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).	
		CC -1- SUBUNIT: MONOMER (BY SIMILARITY).	
		CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).	
		CC -1- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.	

		This SWISS-PROT entry is copyright. It is produced through a collaboration	
		between the Swiss Institute of Bioinformatics and the EMBL Outstation	
		the European Bioinformatics Institute. There are no restrictions on its	
		use by non-profit institutions as long as its content is in no way	
		modified and this statement is not removed. Usage by and for commercial	
		entities requires a license agreement (see http://www.isb-sib.ch/announce/	
		or send an email to license@isb-sib.ch).	

		EMBL; AE006404; AAK05842.1; -	
		DR InterPro; IPR001986; EPSP_synthase.	
		DR Pfam; PF00275; EPSP_synthase; 1.	
		DR ProDom; PD001867; EPSP_synthase; 1.	
		DR PROSITE; PS00104; EPSP_SYNTHASE_1; 1.	
		DR PROSITE; PS00885; EPSP_SYNTHASE_2; 1.	
		Aromatic amino acid biosynthesis; Transferase; Complete proteome.	
		SEQUENCE 430 AA; 45958 MW; CB216F07AA4EE799 CRC64;	

		Query Match 35.2%; Score 806.5; DB 1; Length 430;	
		Best Local Similarity 41.9%; Pred. No. 3e-43;	
		Matches 184; Conservative 77; Mismatches 157; Indels 21; Gaps 7;	

		QY 15 SSGLSGTVRIPGDKSIHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMCARIRKEG 74	
		DB 7 SQGLGRKLKVPVGDKSIHRSIMFGSTAKGTIVYDILRGEDVLSTIEAFRAMGVLELDKG 66	
		QY 75 DTWIDGVNGGILLAEAPLDFGNATGCRMLMGVLGVYDFDSFTFGDASLTAKRPMGRYL 134	
		DB 67 EVITVHGKGISELKAPEKALDMGNSCTSTRLLSIGIAGLPFFETLGDDSLSKRPMDRV 126	
		QY 135 NPUREMGVVKSP--DGDRLPVTLRGPKTPTTYRVPMASAOVKSASVLAGLNTPGITTIV 193	
		DB 127 TPLQLMGAETGGTDKVKLPMTIKGSTHLKAIDYLVPSAVSAOVKSIVFAALQAEGTLTKV 186	
		QY 194 IEPIPTRDHTEKMQOGGANLTVETDADGVTRIRLEGRCKLTGOVIDVDPDSSSTAFLPV 253	
		DB 187 VEKEKTRSRHTSELMVQFGELKYSD-----KTLTPVGQKRLVKGVKVPDVISSAAFV 241	

CC 76 TWIDGVNGGLLAPLDFGNATGRLTMGLGVYDFDSTFGIDASLTKRPMGRVLN 135
 CC 72 IVTIRGVGLGLQPPKAPLNMONSGTSMRLLAGILAAQRFESVLCDESLKRPQRIT 131
 CC 136 PLREMGVQVKSDDRLPVLTRGPKPTPTTYRVPMSAQVKSAVLAGLNTGPIITVIE 195
 CC 132 PLVQMGAKIVSHNFTAPLHISG-RLPTGIDVALPLPSAQLKSLCLLAGLADGTTLRHT 190
 CC 196 PLMTDRHTEKMLGGCANLTVETDADGVTRIRLEGRKLTGGVIDVPGDPSSTAFPLVAA 255
 CC 191 CGISRDHTEKMLPFLFGALIEIKK-----QIIVTGGOKLHGCVLDIVGDLRAAFAFMVAA 245
 CC 256 LLVPGSDVTILNVLNMPRTTGLTLQEMGADIEVNPLRAGGEDVADLRVRSSTLKGVT 315
 CC 246 LIAPRAEVIRNVGNPTAAITLLOKMGRIELHQHFRWGEVADIVVYHSHKURGIT 305
 CC 316 VPEDRAPSMIDEPYILAVAAFAEGATVMNGLEELRVKESDRLSAVANGILKMGVDCDEG 375
 CC 306 VAPWETIANAIDELPIFFIAACAEGTTFVGNLSSELRVKESDRLSAQAQNLQTLGVACDVG 365
 CC 376 ETSLVVRGPRDGKGLGNASGAAVATHLDRHRIAMSFVLMLGVSENPTVDDATMTATSF 435
 CC 366 ADFIHYGRSDRQFL----PARVNSFGDHRHRIAMSLAVAGVRAAGELLIDDGAVAAVSMQP 421
 CC 436 FMDLMAGLCAKIELSDTK 453
 CC 422 FRDFAAAGMNVGERDAK 439

RESULT 7
 AROA_STRPN STANDARD; PRT; 427 AA.
 ID AROA_STRPN STANDARD; PRT; 427 AA.
 AC Q9SA00;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
 DE enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
 GN AROA OR SP1371.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RC STRAIN=NCIMB 40794 / 0100993;
 RC MEDLINE=20069365; PubMed=10601870;
 RC Du W., Wallis N.G., Mazzulla M.J., Chalker A.F., Zhang L., Liu W.-S.,
 Kallender H., Payne D.J.;
 RT "Characterization of Streptococcus pneumoniae 5-enolpyruvylshikimate
 RT 3-phosphate synthase and its activation by univalent cations.";
 RL Eur. J. Biochem. 267:222-227(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TIGR4;
 RC MEDLINE=21357209; PubMed=11463916;
 RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
 RA Peterson S., Heidelberg J., Deboy R.T., Haft D.H., Dodson R.J.,
 RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
 RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
 RA Holtzapple E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L.,
 RA McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,
 RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
 RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
 RT "Complete genome sequence of a virulent isolate of Streptococcus
 RT pneumoniae.";
 RL Science 293:498-506(2001).
 CC -!- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
 CC phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
 CC -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
 CC THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
 CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).

CC -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC EMBL; AF169483; AAD45819.1; -;
 DR EMBL; AE007434; AAK75469.1; -;
 DR TIGR; SP1371; -;
 DR InterPro: IPR001986; EPSP_synthase.
 DR Pfam: PF00275; EPSP_synthase; 1.
 DR ProDom: PD001867; EPSP_synthase; 1.
 DR PROSITE: PS00104; EPSP_SYNTHASE_1; 1.
 DR PROSITE: PS00885; EPSP_SYNTHASE_2; 1.
 KW Aromatic amino acid biosynthesis; Transferase; Complete proteome.
 FT CONFLICT 13 13 S -> I (IN REF. 1).
 FT CONFLICT 71 71 I -> V (IN REF. 1).
 FT CONFLICT 201 201 K -> Q (IN REF. 1).
 FT CONFLICT 303 303 G -> C (IN REF. 1).
 SQ SEQUENCE 427 AA; 45766 MW; 45CE6F4D0D1C7B70 CRC64;

Query Match 35.1%; Score 803.5; DB 1; Length 427;
 Best Local Similarity 42.8%; Pred. No. 4.6e-43;
 Matches 184; Conservative 72; Mismatches 155; Indels 19; Gaps 7;

QY 18 LSGTVRIPGDKSISHRSMFEGGLASGETRITGLLEGEDVINTGKMQAMCARIKKEDTW 77
 DB 10 LHGSIKVPDGSISHRSMFEGGLASGETRITGLLEGEDVINTGKMQAMCARIKKEDTW 69
 QY 78 IIDGVNGVGLLAPLDFGNATGRLTMGLGVYDFDSTFGIDASLTKRPMGRVLNPL 137
 DB 70 TIQGVGMAGLKAQNALNMGNSCTSIKLSGLAGADFEVEMFGDSDLSKRPMDRVTLPL 129
 QY 138 REMGVQVKS-DCDRPLVTLRGPKPTPTTYRVPMSAQVKSAVLAGLNTGPIITVIEP 196
 DB 130 KMGVSISGGTERDPLPLRLKTKNLRIHYELPIASQVKSALMFAALQAKGESVIEK 189
 QY 197 IMTRDHTKMLQCGANLTVETDADGVTRIRLEGRKLTGGVIDVPGDPSSTAFPLVAA 256
 DB 190 EYTRNTHEDMLKQFGHLSV----DG-KKITVQGPQKLTGQKVVVPGDISAFAVLVAGL 244
 QY 257 LVPGSDVTILNVLNMPRTTGLTLQEMGADIEV--INPRLAGGEDVADLRVRSSTLKG 314
 DB 245 IAPNSRLVLQNVGINETRTGIIDVIRAMGKLEITEIDPVAKS---ATLIVESSDLKGT 300
 QY 315 TVPEDRAPSMIDEPYILAVAAFAEGATVMNGLEELRVKESDRLSAVANGILKMGVDC 374
 DB 301 EIGGALIPRLIDELPIALLATQAGVTVIKDAEELKVKETRIQVVVADALNSMGADITP 360
 QY 375 GETSLVVRGPRDGKGLGNASGAAVATHLDRHRIAMSFVLVMG-LVSENPVTVDDATMTATSF 433
 DB 361 TADGMIIKGA-----SALHGARNVTGDRHRCMTAIAALLVADGEVELDRAEINTSY 414
 QY 434 PEFMDLMAGL 443
 DB 415 PSFFDDLESL 424

RESULT 8
 AROA_LACLC STANDARD; PRT; 430 AA.
 ID AROA_LACLC STANDARD; PRT; 430 AA.
 AC P43905;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
 DE enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
 GN AROA.

Lactococcus lactis (subsp. cremoris) (Streptococcus cremoris).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Lactococcus.
ON NCBI_TaxID=1359;
OX [1]
RR SEQUENCE FROM N.A.
RP STRAIN=WG1363 / FJ5876;
RX MEDLINE=95124293; PubMed=7823907;
RA Griffin H.G., Gasson M.J.:
RT "Genetic aspects of aromatic amino acid biosynthesis in Lactococcus
RT lactis.";
RL Mol. Gen. Genet. 246:119-127(1995).
RM -|- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
CC phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
CC THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC -|- SUBUNIT: MONOMER (BY SIMILARITY).
CC -|- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -|- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation at
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).CC -----
CC EMBL; X784113; CAAS5180.1; ..
DR InterPro; IPR001986; EPSP_synthase.
DR Pfam; PF00275; EPSP_synthase; 1.
DR PRODOM; PD001867; EPSP_synthase; 1.
DR PROSITE; PS00104; EPSP_SYNTHASE_1; 1.
DR PROSITE; PS00885; EPSP_SYNTHASE_2; 1.
DR Aromatic amino acid biosynthesis; Transferase.
SQ SEQUENCE 430 AA; 45804 MW; C5A197A9072C9D5 CRC64;

Query Match 35.1%; Score 802.5; DB 1; Length 430;
Best Local Similarity 41.2%; Pred. No. 5.4e-43;
Matches 181; Conservative 79; Mismatches 158; Indels 21; Gaps 7;

Qy 15 SSGLSGVTRIPGDCKSIHRSFMFGGLASGETRITCLLEGEDVINTGKAMQAGRIKEG 74
Db 7 SQGLKGRUGLPVDCKSIHRSFMFGGLASGETRITCLLEGEDVINTGKAMQAGRIKEG 66

Qy 75 DTWIIDGVNGGLLAPEAPLFQGNATGCRLTGMVLGYDFDSTFIGNASLTKRPMGRVL 134
Db 67 QVITVHGOGISKLKEPEKALDNMGNSGTSTRLLSGILAGLPFEATLFGDDSLSKRPMDRV 126

Qy 135 NPLREMGVOVKSE-DGDRLLPTVLRGPKTPPTYRYVPMSAQVKSALLAGLNTGGITTV 193
Db 127 TPLQMGAIEIVGQTDKVLPMTIGKSAHLKAIDYILPVASQVKSAVIFAALQAEGTLKV 186

Qy 194 TEPIWTRDHTTEKMLOGFGANLTVE TDAGVTRIRLEGRKGKTGVQIDVPDPSSSTAFLPV 253
Db 187 VBEKTRTSHEMLVVOFGIEITVSD-----KTLVPGGOKLLGQEVTPDGISSAAFWLV 241

Qy 254 AALLVPGSDVTILNLMNPTRTGILLTLQEMGADIEVINPRLAGGED----VADLRVRSS 309
Db 242 AGLVVNSGLILENVGINETRTGILEVIQAMGGQLEILE-----QDEVAKAATLKVKAS 295

Qy 310 TLKGVTVPEDRAPSMIDEYPILAAVAFAEAGATVNGMLEELRVKESDRLSAVANGKLKG 369
Db 296 QLKGTEISGDLPLRLDEIPIALTALTAQAEKGTIIINDAAELKKVETRIDIAVVADALNSMG 355

Qy 370 VPDCEGETSLVYGRPDGKGLGNASGAATAHLDRHIAAMSFLVMGLVSEN-PVTYDDATM 428
Db 356 ANIEPTDDGMIIQG---GTKL-HAPENAITLGDRHIGMMVAIALLVENGEEILERAEA 411

Qy *429 IATSPFEFMDLMAGLGAKI 447
Db 412 IQTSYPSFFDLEKLSGNL 430


```

Db 239 ACAGMPNSRIVLKNVGLNPTGIGTIDVQLQNGAKLEIKPSADSGAEPYGDLLIETSSLK 298
QY 313 GVTVPEDRAPSMIDEPYILAVAAFAEGATVMNGLEELRVKESDRLSAVANGKLNGVDC 372
Db 299 AVEIGGDIIPRLIDEPIIALLATQAEETTVIKDAAELKVKETNRIDTVVSELKLGAEI 358
QY 373 DEGETSLVVRGPDGKGLGNASGAAVATHLDRHIAFLVGLVSENPTVDDATMIATS 432
Db 359 EPTADGMKYVGKOTLKG-----GAAVSSGHDHRIGMGLGIASCITEPIETIEHTDAIHS 413
QY 433 FPEFMDLMAGLGAK 446
Db 414 YTFPEHLNKLKSK 427

RESULT 10
AROA_AQUAE STANDARD; PRT; 431 AA.
O67494;
30-MAY-2000 (Rel. 39, Created)
30-MAY-2000 (Rel. 39, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GN AROA OR AQ.1536.
OS Aquifex aeolicus.
OC Bacteria; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
Graham D.E., Overbeek R., Snead M.A., Keller M., Aufay M., Huber R.,
Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
aeolicus.";
RL Nature 392:353-358(1998).
CC -!- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; AE000744; AAC07443.1; -
DR InterPro; IPR001986; EPSP_synthase.
DR Pfam; PF00275; EPSP_synthase; 1.
DR PROSITE; PD001867; EPSP_synthase; 1.
DR PROSITE; PS00104; EPSP_synthase; 1.
DR PROSITE; PS00885; EPSP_synthase; 2; 1.
KW Aromatic amino acid biosynthesis; Transferase; Complete proteome.
SQ SEQUENCE 431 AA; 47793 MW; EF842512EBE41D2A CRC64;

Query Match 33.5%; Score 766.5; DB 1; Length 431;
Best Local Similarity 40.7%; Pred. No. 9.2e-41;
Matches 174; Conservative 73; Mismatches 169; Indels 11; Gaps 6;

QY 14 KSGLSGTVRPGDKSISHRSMFVGGLASGETRITGLLEGEDVINTGKAMQANGAIRKE 73
Db 6 KIKRVKGLRVPDSKSIHTRAFTLIGALASGETLVRKPLISGDTLATLILKAIRTKVREG 65

```

```

QY 74 GDTWIIDGVNGGILLAPEDFGNATGCRLTWGLVGVYDFDSTFTIGDASLTFRPMGRV 133
Db 66 KEVLLIEG-RNYTFLEPHDVLDAKNSGTTARIMSGVLSTQPFPSVLTGDSLNRPMLRV 124
QY 134 LNPRLREMGVQVKS-EDGDRLPVTLRGPKTPTITYRVPMAAQVKSAVLLAGLNTPGITT 192
Db 125 VEPRLREMGAKIDRGEGNKLPPIAIRGNL-KGISYFNKKSAAQVKSALLLAGLRAEGMTE 183
QY 193 VIEPIMTROHTEKMLQGFGANLTVETDADGVRTIRLEGRKGLTGQVIVDPGDSSTAFPL 252
Db 184 VVEPYLSRDRHTEKMLQGFANLTVETDADGVRTIRLEGRKGLTGQVIVDPGDSSTAFPL 242
QY 253 VAALLVPGSDVTILNVLNPNTRTGLTILTOEMGADIEVINPRLAGGEDVADLVRVS-STL 311
Db 243 ALATLAPEGEIRKEKLELLNPNTRDGFYRKLEMGDLSFENYRELSNPNMADLVVRVDNL 302
QY 312 KGVTVPEDRAPSMIDEPYILAVAAFAEGATVMNGLEELRVKESDRLSAVANGKLNGVD 371
Db 303 KPVKVSPEEVPITLIDEIPILAVLMAFADGVSEVKAGELRYKESDRILKAIVTLNRKLG 362
QY 372 CDEGETSLVVRGPDGKGLGNASGAAVATHLDRHIAFLVGLVSENPTVDDATMIAT 431
Db 363 VEEFEDGFALHGTKEIK-----GGVETFKDHRIAMAFVGLVVEEVIIDHPECVTY 416
QY 432 SFPEFMD 438
Db 417 SYPEFWE 423

RESULT 11
AROA_THEME STANDARD; PRT; 410 AA.
AC Q9WY10;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GN AROA OR TM0345.
OS Thermotoga maritima.
OC Bacteria; Thermotogales; Thermotoga.
OX NCBI_TaxID=2336;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
genome sequence of Thermotoga maritima.";
RL Nature 399:323-329(1999).
CC -!- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; AF001715; AAD35431.1; -
DR TIGR; TW0345; -

```

RT reveals hypervariable sequences. ;

RL Nature 403:665-668(2000).

CC -1- CATALYTIC ACTIVITY: phosphoenolpyruvate + 3-phosphoshikimate -

CC phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.

CC ProDom: P0001867; EPSP_synthase; 1.

CC PROSITE: PS00104; EPSP_SYNTHASE_1; 1.

CC PROSITE: PS00885; EPSP_SYNTHASE_2; 1.

CC Aromatic amino acid biosynthesis; transferase; Complete proteome.

CC Aromatic amino acid biosynthesis; transferase; Complete proteome.

CC -1- SUBUNIT: MONOMER (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).

CC -1- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>

CC or send an email to license@isb-sib.ch).

CC

CC EMBL: X89371; CAA61554.1; -.

CC EMBL: AL139076; CAB73153.1; -.

CC InterPro: IPR001986; EPSP_synthase.

CC Pfam: PF00275; EPSP_synthase; 1.

CC ProDom: P0001867; EPSP_synthase; 1.

CC PROSITE: PS00104; EPSP_SYNTHASE_1; 1.

CC PROSITE: PS00885; EPSP_SYNTHASE_2; 1.

CC Aromatic amino acid biosynthesis; transferase; Complete proteome.

CC CONFLICT 152 152 K -> N (IN REF. 1).

CC CONFLICT 158 158 F -> Y (IN REF. 1).

CC CONFLICT 179 179 D -> N (IN REF. 1).

CC CONFLICT 183 183 T -> A (IN REF. 1).

CC CONFLICT 217 217 N -> S (IN REF. 1).

CC CONFLICT 241 241 V -> A (IN REF. 1).

CC SEQUENCE 428 AA; 4728 MW; 465E2B50F39FC5AB CRC64;

Query Match 25.8%; Score 590; DB 1; Length 428;

Best Local Similarity 34.0%; Pred. No. 8.2e-30;

Matches 146; Conservative 94; Mismatches 167; Indels 22; Gaps 11;

QY 24 IPGDKSISHRSMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTWIDGVG 83

DB 17 IAADKSISHRFAIFSLTQENKAQNYLLAQTDLNTLEIKNLGAKIKQKDS--CVKIIP 74

QY 84 NGGLAPEAPLDFGNAATGCRITMG-LVGVDYDFDSTFI--GDASLTGKRMGRVNLPLREM 140

DB 75 PKELSPNCLDCGSGTAMRLMIGLAGISGF---FVLGDKYLNRRPMRRISKPLTQI 131

QY 141 GVQKSDGSDGL--PVTLRGKPTPTTYRVPMAQSAQVAVLLAGNLTPGTTVIEPIWT 199

DB 132 GARIYGRNEANLAPLCIEGOKLA--ENFKSEISSAQVKTAMLSAFRADNVCTFSEISL 190

QY 200 RDHTEKMGQFGANLVETDADGVRTIRLEGKCKLTGQVIDVPDPSPTAFPLVAAILVP 259

DB 191 RNHSENMLKAMKAPIRVSDGLSLEINPL--KKPLAKNIILPDPSSAFYFLVAAILP 248

QY 260 GSDVTTILNVLNPTRTGLITLQEMGADIEVINPRLAGGEDVADLRVRSSTLKGVTVPD 319

DB 249 KQIILKNILLNPTRIEAYKILQMGAKLEMTITQ--NDFETIGEIRVESKLVGIEV-KD 306

QY 320 RAPSMDIEYPIAVAAFAEGATVMNGLELRYKESDRLSAVANGKLKGVDCDEGTSU 379

DB 307 NIAMLIDEAPALATAFALAKGKSSLNAKELRVKESDRIVAVVWENKLCGVARELDDGF 366

QY 380 VVRGRPDGKGLGNAGAAVATHLDHRIAMSFVLGVSNPVTVDDATMIATSFPEFMDL 439

DB 367 EIEGCEK-----SSKIKSYGDHRIAMSFVAILGLLC--GIEIDSDCIKTSFPNFIEI 418

QY 440 MAGLGAKIE 448

DB 419 LSNLGARDI 427

RESULT 13

InterPro: IPR001986; EPSP_synthase.

Pfam: PF00275; EPSP_synthase; 1.

ProDom: P0001867; EPSP_synthase; 1.

PROSITE: PS00104; EPSP_SYNTHASE_1; 1.

PROSITE: PS00885; EPSP_SYNTHASE_2; 1.

Aromatic amino acid biosynthesis; transferase; Complete proteome.

SEQUENCE 410 AA; 45341 MW; DC9F7D44792CA69F CRC64;

Query Match 30.4%; Score 695; DB 1; Length 410;

Best Local Similarity 40.5%; Pred. No. 2.4e-36;

Matches 170; Conservative 67; Mismatches 161; Indels 22; Gaps 10;

QY 24 IPGDKSISHRSMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTWIDGVG 83

DB 4 VPPDKSITHRALISALAEESTLYNLRLCJDTERTHDILEKGTFR--EGD-WEKMKVF 60

QY 84 NGGLAPEAPLDFGNAATGCRITMGLGVGYDFDSTFTGDSLTGKRMGRVNLPLREMGVQ 143

DB 61 PKPFAEPIELPCGNSGTTIRLMSGVLASYEFVILGCDPSLSRRPMRRVIEPLEMMGAR 120

QY 144 VKSEGDRLPVTLRGKPTPTTYRVPMAQSAQVAVLLAGNLTPGTTVIEPIWTRDHT 203

DB 121 FMAQQNYLPAIMKGNHL--SGISYKTSIAVSAQVAVLLAGNLASRTIVIEPAKSRDHT 179

QY 204 ERMQLGFGANLVETDADGVRTIRLEGKCKLTGQVIDVPDPSPTAFPLVAAILVPGSDV 263

DB 180 ERMKL-----NLGVPEVEGTRVV--LE-PATFRGFTMKVPGDISAAEFVLGAHPNARI 233

QY 264 TILNVLNPTRTGLITLQEMGADI--EVINPRLAGGEDVADLRVRS--TLKGVTVPEDR 320

DB 234 TVTVDGLNPTRTGLLEVMKLMGANLEWETENL---EPICGTVRVTSPNLKGVVVVPEHL 290

QY 321 APSKIDIEYPIAVAAFAEGATVMNGLELRYKESDRLSAVANGKLKGVDCDEGETSLV 380

DB 291 VPLMIDELPLVALLGVFAEGETVVRNABELRKESDRIRVLVENFKRLGVIEEFKDGFK 350

QY 381 VVRGRPDGKGLGNAGAAVATHLDHRIAMSFVLGVSNPVTVDDATMIATSFPEFMDL 440

DB 351 IVCKQSTK-----GGSVDPEGDRHMAMLFSTAGLVSEEGVDVKDHECVAVSFPNFYELL 404

RESULT 12

ARO_A_CAMJE STANDARD; PRT: 428 AA.

AC P52312: Q9pp36;

DT 01-OCT-1996 (Rel. 34, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DE 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-

EN encolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).

GN AROA OR CJO895C.

OS Campylobacter jejuni.

OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;

OC Campylobacter.

OX NCBI_TaxID=197;

QY [1]

SEQUENCE FROM N.A.

RC STRAIN=81116;

RX MEDLINE=97128776; PubMed=8973316;

RA Woesten M.M.S.M., Dubink V.H.J., van der Zeijst B.A.M.;

RT "The ara gene of Campylobacter jejuni.";

RL Gene 181:109-112(1996).

RP [2]

SEQUENCE FROM N.A.

RC STRAIN=NCTC 111168;

RX MEDLINE=20150912; PubMed=10688204;

RA Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,

RA Braham D., Chillingworth T., Davies R.M., Feltwell T., Holroyd S.,

RA Jagels K., Karlyshev A.V., Moutie S., Fallan M.J., Penn C.W.,

RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,

RA Whitehead S., Barrall B.G.;

RT "The genome sequence of the food-borne pathogen Campylobacter jejuni

```

ARO_ASTAAU          STANDARD;          PRT;          430 AA.
ID ARO_ASTAAU          AC Q05615;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GN AROA.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCCT 8325-4;
RC MEDLINE=93381456; PubMed=8371108;
RA O'Connell C.M., Pattee P., Foster T.J.;
RA "Sequence and mapping of the aroA gene of Staphylococcus aureus
8325-4.";
J. Gen. Microbiol. 139:1449-1460(1993).
CC -1- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC -1- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
-----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL: L05004; AAA1897.1; -.
CC InterPro: IPR001986; EPSP_synthase.
CC Pfam: PF00275; EPSP_synthase; 1.
CC ProDom: PD001867; EPSP_synthase; 1.
CC PROSITE: PS00104; EPSP_SYNTHASE.1; 1.
CC PROSITE: PS00885; EPSP_SYNTHASE.2; 1.
CC Aromatic amino acid biosynthesis; transferase.
KW AROMATIC AMINO ACID BIOSYNTHESIS; TRANSFERASE.
SQ SEQUENCE 430 AA; 46852 MW; E5FC878EALC23C20 CRC64;
-----
Query Match 25.2%; Score 577; DB 1; Length 430;
Best Local Similarity 33.6%; Pred. No. 5.3e-29;
Matches 144; Conservative 82; Mismatches 179; Indels 24; Gaps 8;

QY 15 SSGLSGTVRIPGDKSIGHSRFGGLASGETRIRGLLEGEDVINTGRAMQAMGARIRKE 74
   | | | : : : | | | | : | | | | | | | | | | | | | | | | | | |
Db 10 SGPLKGEIEVPGDKSWTHRAIMLASLAEGYSTIYKPLLGDCRTMDIFRHLGVEIKEDD 69
   : : : : : | | | | | | | | | | | | | | | | | | | | | | | |
QY 75 DTWIIDGVNGGLLAPLEAPLDFGNAATGCRMTGLMGVGYDFDSTFIDGASLTKRPMGRVL 134
   : : : : : | | | | | | | | | | | | | | | | | | | | | | | |
Db 70 EKLVTSPGY-QVNTPHQVLYTNGSGTTRLLAGLLSGLSGNSVLSGDVSIGKRPMDRVL 128
   : : : : : | | | | | | | | | | | | | | | | | | | | | | | |
QY 135 NPLREMGVQKSEGDRLPYTLRGPKPTPTTYRVPMAAQVKSAVLAGNLTCGTTVI 194
   | | : : : : : | | : : : : : | | : : : : : | | : : : : : |
Db 129 RPLKMDANTEGEDNTYPLIHK-PSVIKGINYOMEVASAQVKSAILFASLFSKREPTIK 187
   | | : : : : : | | : : : : : | | : : : : : | | : : : : : |
QY 195 EPIMTRDQETKMLQGF-----GANLTVETADAGVYTRLEGRGLKTGVIDVPGDPSSTA 249
   | : : | | | | | | | | : : : : : | | | | | | | | | | | | |
Db 188 ELDSRNHETFMFKHFNPIEAEGLSINTTPEAIRYIKPAD-----FHVPGDISAA 239
   | : : | | | | | | | | : : : : : | | | | | | | | | | | | |
QY 250 FPLVAALLVPGSDVTTILNLMNPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRV-S 308
   | : | | | : | | | | | | | | | | | | | | | | | | | | | | :
Db 240 FFIVAALLITPGSDVTHNVGINQTRSGIIDIVKMGNGNIQLFN-QTTGAETASIRIQYT 298
   | : : : : | | | | | | | | | | | | | | | | | | | | | | :
QY 309 STLKGVTVPEDRAPSMIDEYPIIAVAFAAFAGVATVMNGLELRVKESDRLSAVANGKLN 368
   | : : : : | | | | | | | | | | | | | | | | | | | | | | :

```

DR TIGR: HP0401; -
DR InterPro: IPR001986; EPSP_synthase.
DR Pfam: PF00275; EPSP_synthase; 1.
DR ProDom: PD001867; EPSP_synthase; 1.
DR PROSITE: PS00104; EPSP_SYNTHASE_1; 1.
DR PROSITE: PS00885; EPSP_SYNTHASE_2; 1.
KW Aromatic amino acid biosynthesis; Transferase; Complete proteome.
SQ SEQUENCE 429 AA; 47240 MW; 19545753E081FDAE CRC64;

Query Match 22.7%; Score 518.5; DB 1; Length 429;
Best Local Similarity 33.5%; Pred. No. 2.2e-25;
Matches 145; Conservative 74; Mismatches 171; Indels 43; Gaps 13;

QY 27 DKSTSHSFEGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEG-----DTWI 78
DB 10 DKSLSHRAVIFSLAQKPCFVRNFMAGEDCLSSLEIAQNLGAKVENTAKNSFKITPPTI 69
QY 79 IDGVNGGLLAPEAPLDFGNAATGCRMTGLVGVYDFDSTFIFGDASLTGKPMGRVNLPLR 138
DB 70 KE-----PNKILNCNNSGTTMRLYSGLLSAQKGLFVLSGDSNLNARPMKRIEPLK 120
QY 139 EMGVQVKS-EDGDRLPVTLRGPKTPIT---YRPNASAQVKSAVLLAGLNTPGITTVI 194
DB 121 AFGAKILGREDNHFAPLVILG-----SPLKACHYESPIASAQVKSFAFSLAQAGASTYK 176
QY 195 EPIMTDRHTEKMLQGFANLTVETDADGVRTIR-LEGRKLTGQVIDVPGDPSSTAPPLV 253
DB 177 ESELSRNHTEIMLSLGADI---HNQDGVLKISPLE--KPLEAFDETIANDPSSAFFAL 231
QY 254 AALVPGSDVTILNLMNPTTGLILLOENGADIE--VINPRLAGGEDVADLRVSTL 311
DB 232 ACAITPKSRLLKNVLLNPTRIEAFVLLKMGASIEYAIOSKDL---EMIGDIIVEHAPL 288
QY 312 KGVTVPEDRAPSMIDEXPIAVAAFAEGATVMNGLELRYKESDRLSAVANGKLVGD 371
DB 289 KAINIDQINIA-SLIDEIPALSIAMLFAGKSKMKNANDLRAKESDRKAVVSNFKALGIE 347
QY 372 CDEGETSLVVRGPDGKGL----GNASGAAVATHLDHRIAMSFVLMGLVSENPVVDAT 427
DB 348 CEEFEDGFYVEGLEDISPLKQRFSGRIKPLIKSFNDHRIAMSFVLTIAL--PLEIDNLE 405
QY 428 MIATSPFEEDIM 440
DB 406 CANISFPQFKHLL 418

Search completed: August 15, 2002, 14:06:06
Job time: 405 sec

QY 139 EMGVQVKS-EDGDRLPVTLRGPKTPITITYRVPMSAQVKSAVLLAGLNTPGITTVIEPI 197
DB 121 AFGAKILGREDNHFAPLVILGCPKA-CDYESPIASAQVKSFAFSLAQAGISAYKESE 179
QY 198 MTRDHTKMLQGFANLTVETDADGVRTIR-LEGRKLTGQVIDVPGDPSSTAPPLVAAL 256
DB 180 LSRNHTIMLSLGANI---QNQDGVLKISPLE--KPLESFDETIANDPSSAFFALACA 234
QY 257 LVPGSDVTILNLMNPTTGLILLOENGADIE--VINPRLAGGEDVADLRVRSSTLKGVT 315
DB 235 ITPKSRLLKNVLLNPTRIEAFVLLKMGAEIYVIOSK--DLEVIGDIYIEHAPLKAIS 292
QY 316 VPEDRAPSMIDEXPIAVAAFAEGATVMNGLEELRYKESDRLSAVANGKLVGDDEG 375
DB 293 IDQINIA-SLIDEIPALSIAMLFAGKSKMKNANDLRAKESDRKAVVSNFKALGIECEE 351
QY 376 ETSLVVRGPDGKGLGNAS-----GAAVATHLDHRIAMSFVLMGLVSENPVVD 425
DB 352 EDGFIY-----EGLDASQLKQHSKIKPPIIKSFNDHRIAMSFVLTIAL--PLEIDN 403
426 ATMIATSPFE 436
404 LECANISPTTF 414

RESULT 15
ID AROA_HELPY STANDARD; PRT; 429 AA.
AC P56197;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GN AROA OR HP0401.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter;
OX NCBI_Taxid=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Kervavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RA "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori.";
RL Nature 388:539-547(1997).
CC -1- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
CC phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC -1- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
CC THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE000556; AAD07470.1; -

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 15, 2002, 13:58:51 ; Search time 51.67 Seconds
(without alignments)
1523.373 Million cell updates/sec

Title: US-09-464-099A-3
Perfect score: 2288
Sequence: 1 MSHGASSRPATARKSSGLSG.....FMDLMAGLAKIELSDTKAA 455

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- SPTREMBL_19:*
- 1: sp_archaea:*
 - 2: sp_bacteria:*
 - 3: sp_fungi:*
 - 4: sp_human:*
 - 5: sp_invertebrate:*
 - 6: sp_mammal:*
 - 7: sp_mhc:*
 - 8: sp_organelle:*
 - 9: sp_phage:*
 - 10: sp_plant:*
 - 11: sp_rodent:*
 - 12: sp_virus:*
 - 13: sp_vertebrate:*
 - 14: sp_unclassified:*
 - 15: sp_rvirus:*
 - 16: sp_bacteriap:*
 - 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2104	92.0	455	16 Q92SV5	Q92sv5 rhizobium m
2	1861.5	81.4	480	2 Q9AGV2	Q9agv2 brucella ab
3	1789	78.2	452	16 Q98CC1	Q98cc1 rhizobium l
4	1133	49.5	443	16 Q9A2H2	Q9a2h2 caulobacter
5	989.5	43.2	746	16 Q9H269	Q9hz69 pseudomonas
6	912.5	39.9	454	16 Q9PB21	Q9pb21 xylella fas
7	884	38.6	431	16 Q9KCA6	Q9kca6 bacillus ha
8	820	35.8	430	16 Q99ZB3	Q99zb3 streptococc
9	815.5	35.6	428	16 Q9A2A5	Q9a2a5 listeria in
10	781.5	34.2	428	2 Q9ANY6	Q9any6 enterococcu
11	610	26.7	432	16 Q99U25	Q99u25 staphylococ
12	470.5	20.6	207	2 Q9RHZ8	Q9rhz8 pseudomonas
13	470.5	20.6	439	17 Q9HQCL	Q9hqcl halobacteri
14	416.5	18.2	440	2 Q9L213	Q9l213 streptomyce
15	401.5	17.5	426	16 Q9KR80	Q9kr80 vibrio chol
16	378.5	16.5	427	2 Q93ED4	Q93ed4 yersinia ru

17	351.5	15.4	428	16 Q97KM2	Q97km2 clostridium
18	332.5	14.5	438	2 Q9K4A7	Q9k4a7 streptomyce
19	324.5	14.2	447	16 Q9K9D5	Q9k9d5 bacillus ha
20	306	13.4	516	10 Q946V0	Q946v0 dicliptera
21	299.5	13.1	511	10 Q93VK6	Q93vk6 oryza sativ
22	299	13.1	444	10 Q24566	Q24566 zea mays (m
23	298	13.0	516	10 Q946U9	Q946u9 dicliptera
24	296	12.9	521	10 Q9FVP6	Q9fvp6 arabiidopsis
25	293.5	12.8	445	8 Q95AK0	Q95ak0 eleusine in
26	292.5	12.8	445	8 Q95AK1	Q95ak1 eleusine in
27	285	12.5	433	16 Q9JTT3	Q9jtt3 neisseria m
28	280	12.2	414	17 Q980I5	Q980i5 sulfolobus
29	275	12.0	433	16 Q9JYU1	Q9jyu1 neisseria m
30	261	11.4	408	17 Q96Y91	Q96y91 sulfolobus
31	260.5	11.4	427	17 Q9YC47	Q9yc47 aeropyrum p
32	229	10.0	332	2 Q9ZEQ0	Q9zeq0 actinobacil
33	223	9.7	410	17 Q978S3	Q978s3 thermoplasm
34	216.5	9.5	391	10 Q80428	Q80428 oryza sativ
35	204.5	8.9	448	2 Q9L1U5	Q9l1u5 streptomyce
36	201.5	8.8	446	2 Q9SON1	Q9son1 streptomyce
37	185.5	8.1	419	16 Q99Z78	Q99z78 streptococc
38	184.5	8.1	417	16 Q9K1Q9	Q9k1q9 neisseria m
39	184	8.0	418	16 Q97DD9	Q97dd9 clostridium
40	183.5	8.0	417	16 Q9JWS7	Q9jws7 neisseria m
41	182	8.0	428	16 Q9K6E5	Q9k6e5 bacillus ha
42	181	7.9	419	2 Q9SSWS	Q9ssws escherichia
43	179	7.8	347	10 Q9AT37	Q9at37 lolium rigi
44	179	7.8	423	16 Q927U1	Q927u1 listeria in
45	178.5	7.8	417	2 Q9EXE3	Q9exe3 mycobacteri

ALIGNMENTS

RESULT 1

Q92SV5 ID Q92SV5 PRELIMINARY; PRT; 455 AA.
AC Q92SV5;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PUTATIVE 3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE PROTEIN (EC 2.5.1.19).
OS Rhizobium melloti (Sinorhizobium melloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
ON NCBI_TaxID=382;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21368234; PubMed=11474104;
RA Galibert F., Finan T.M., Long S.R., Puehler A., Abola P., Ampe F., Barloy-Hubier F., Barnett M.J., Becker A., Boistard P., Bothe G., Boutry M., Bowser L., Buhrmester J., Cadieu E., Capela D., Chain P., Cowie A., Davis R.W., Dreano S., Federspiel N.A., Fisher R.F., Gloux S., Godrie T., Goffeau A., Golding B., Guzy J., Gurjal M., Hernandez-Lucas I., Hong A., Huizar L., Hyman R.W., Jones T., Kahn D., Kahn M.L., Kalman S., Keating D.H., Kiss E., Komp C., Leleure V., Masuy D., Palm C., Peck M.C., Pohl T.M., Portetelle D., Purnelle B., Ramsperger U., Surzycki R., Thebault P., Vandenbol M., Vorholter F.J., Weidner S., Wells D.H., Wong K., Yeh K.-C., Batut J.;
RT "The composite genome of the legume symbiont Sinorhizobium melloti.";
RL Science 293:668-672(2001).
DR EMBL; AL591783; CAC41690.1; -;
KW Transferase; Complete proteome.
SQ SEQUENCE 455 AA; 47900 MW; 97659E1C7E1021B5 CRC64;

Query Match 92.0%; Score 2104; DB 16; Length 455;
Best Local Similarity 90.8%; Pred. No. 1.2e-119;
Matches 413; Conservative 18; Mismatches 24; Indels 0; Gaps 0;

QY 1 MSHGASSRPATARKSSGLSGTVPRIKSGTSHRSFMEGLASGETRITGLLEGEDVINTG 60

```
Db 1 MSHGSSNPATARKSSDGLKGLRIIPGDKSISHSRSMFGLAAGETRIITGLGEDVINTG 60
QY 61 KAMQAMGARIRKEGDTWIIDGVNGGLLAPEAPLDFGNAATGCRLTGMLGVYDFDSTFI 120
Db 61 KAMQAMGARIRKEGDTWIIDGVNGGALLAPEAPLDFGNAGTGCRLTGLGVYDFDSTFI 120
QY 121 GDSLTKRPMGRVLPNPLREMGVQVKSSEGDRPLVTLRGPKTPPTTYRVPMSAQVKSAY 180
Db 121 GDSLTKRPMGRVLPNPLREMGVQVKSSEGDRPLVTLRGPKTPPTTYRVPMSAQVKSAY 180
QY 181 LLAGLNTPGITTVIEPIIMTRDHTKMLQGFANLVETDADGVRIIRLEGKGLTGQVID 240
Db 181 LLAGLNTPGITTVIEPIIMTRDHTKMLQGFANLVETDADGVRIIRLEGKGLTGQVID 240
QY 241 VPGDPSSTAFPLVALLVPGSDVTILNLMNPTRTGLILTQEMGADIEVINPRLAGGED 300
Db 241 VPGDPSSTAFPLVALLVPGSDVTILNLMNPTRTGLILTQEMGADIEVINPRLAGGED 300
QY 301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPIILAVAAFAEGATVMNGLELRVKESDRLSA 360
Db 301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPIILAVAAFAEGATVMNGLELRVKESDRLSA 360
QY 361 VANGKLNGVDCDEGETSLVVRGPRDGGKLGNAAGAAVATHLDHRIAMSFVLMGLVSENP 420
Db 361 VADGLKNGVDCDEGASLVVRGPRDGGKLGKISGQVKTGLDHRHIAFVLMGLASEHP 420
QY 421 VTVDATMIATSPFPEMDLMAGLAKIELSDTKAA 455
Db 421 VTVDATMIATSPFPEMDLMAGLAKIELSDTKAA 455
```

```
RESULT 2
Q9AGV2 ID Q9AGV2 PRELIMINARY; PRT; 480 AA.
AC Q9AGV2;
DT 01-JUN-2001 (TremBLrel. 17, Created)
DT 01-JUN-2001 (TremBLrel. 17, Last sequence update)
DT 01-OCT-2001 (TremBLrel. 18, Last annotation update)
DE 5-ENOLPYRUVIL SHIKIMATE 3-PHOSPHATE SYNTHASE (EC 2.5.1.19).
GN AROA.
OS Brucella abortus.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=235;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2308;
RA Gan T., Essenberg R.C.;
RT "Characterization of the aroA gene of Brucella abortus and construction of an aroA mutant.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF326475; AAK27445.1; -.
DR InterPro; IPR001986; EPSP_synthase.
DR Pfam; PF00275; EPSP_synthase; 1.
DR ProDom; PD001867; EPSP_synthase; 1.
DR PROSITE; PS00104; EPSP_SYNTHASE_1; UNKNOWN_1.
DR PROSITE; PS00885; EPSP_SYNTHASE_2; UNKNOWN_1.
KW Transferase.
SQ SEQUENCE 480 AA; 50634 MW; 3D55323944DA8C91 CRC64;
```

```
Query Match 81.4%; Score 1861.5; DB 2; Length 480;
Best Local Similarity 81.7%; Pred. No. 5.9e-105;
Matches 365; Conservative 35; Mismatches 44; Indels 3; Gaps 1;

QY 1 MSHGASSRPARATARKSSGLSTGTVRIPGDKSISHSRSMFGLASGETRITGLGEDVINTG 60
Db 31 MSHSACPATARKSQAALTGEIRIPGDKSISHSRSMFGLASGKTRITGLGEDVINTG 90
QY 61 KAMQAMGARIRKEGDTWIIDGVNGGLLAPEAPLDFGNAATGCRLTGMLGVYDFDSTFI 120
Db 91 KAMQAMGARIRKEGDTWIIDGVNGGCLLQPEAPLDFGNAGTGARLTGMLGVYDMKTSFI 150
```

```
QY 121 GDSLTKRPMGRVLPNPLREMGVQVKSSEGDRPLVTLRGPKTPPTTYRVPMSAQVKSAY 180
Db 151 GDSLSKRPMGRVLPNPLREMGVQVKAEGDRPLTLIGPRTANPIAYRVPMSAQVKSAY 210
QY 181 LLAGLNTPGITTVIEPIIMTRDHTKMLQGFANLVETDADGVRIIRLEGKGLTGQVID 240
Db 211 LLAGLNTPGITTVIEPIIMTRDHTKMLQGFADLTVDKGVRIIRIVGGKLTGTID 270
QY 241 VPGDPSSTAFPLVALLVPGSDVTILNLMNPTRTGLILTQEMGADIEVINPRLAGGED 300
Db 271 VPGDPSSTAFPLVALLVPGSEVTIRNVMNPTRTGLILTQEMGADIEIDPRLAGGED 330
QY 301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPIILAVAAFAEGATVMNGLELRVKESDRLSA 360
Db 331 VADLRVRSSTLKGVTVPEDRAPSMIDEYPIILAVAAFAEGATVMNGLELRVKESDRLSA 390
QY 361 VANGKLNGVDCDEGETSLVVRGPRDGGKLGNAAGAAVATHLDHRIAMSFVLMGLVSENP 420
Db 391 VARGLEANGVDCDEGETSLVVRGPRDGGKLG--GGTVATHLDHRIAMSFVLMGLASEKP 447
QY 421 VTVDATMIATSPFPEMDLMAGLAKI 447
Db 448 VTVDSTMIATSPFPEMDLMAGLAKI 474
```

```
RESULT 3
Q98CC1 ID Q98CC1 PRELIMINARY; PRT; 452 AA.
AC Q98CC1;
DT 01-OCT-2001 (TremBLrel. 18, Created)
DT 01-OCT-2001 (TremBLrel. 18, Last sequence update)
DT 01-OCT-2001 (TremBLrel. 18, Last annotation update)
DE 3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE.
GN MLL5213.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Iidesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shlimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL; AP003006; BAB51700.1; -.
DR InterPro; IPR001986; EPSP_synthase.
DR Pfam; PF00275; EPSP_synthase; 1.
DR ProDom; PD001867; EPSP_synthase; 1.
KW Transferase; Complete proteome.
SQ SEQUENCE 452 AA; 47455 MW; 2B52983E3523B938 CRC64;
```

```
Query Match 78.2%; Score 1789; DB 16; Length 452;
Best Local Similarity 78.0%; Pred. No. 1.3e-100;
Matches 352; Conservative 33; Mismatches 62; Indels 4; Gaps 2;

QY 1 MSHGASSRPARATARKSSGLSTGTVRIPGDKSISHSRSMFGLASGETRITGLGEDVINTG 60
Db 1 MSHAAAKPATARKSQAALTGEIRIPGDKSISHSRSMFGLASGETRITGLGEDVINTG 60
QY 61 KAMQAMGARIRKEGDTWIIDGVNGGLLAPEAPLDFGNAATGCRLTGMLGVYDFDSTFI 120
Db 61 KAMQAMGARIRKEGDTWIIDGVNGGCLLQPEAPLDFGNAGTGARLTGMLGVYDMKTSFI 120
QY 121 GDSLTKRPMGRVLPNPLREMGVQV--KSEGDRLPVTLRGPKTPPTTYRVPMSAQVKSAY 179
Db 121 GDSLTKRPMGRVLPNPLREMGVQV--KSEGDRLPVTLRGPKTPPTTYRVPMSAQVKSAY 179
```


Qy	4	GASSRPATARKSSGLSTVRIPGDKSI	SHRSFMFGGLASGETRITGLLEGGEDVINTGKAM	63
Db	5	GLKSAPGGA-----LRGIVRAPGDKSI	SHRSMTLIGALATGTTTVEGLLEGGDVLATARAM	59
Qy	64	QAMGARIRKEG-DTWIIDGVNGGLLAP	ELPEAPLDFGNAATGCRITMGLVGVYDFDSTFTGD	122
Db	60	QATGARIRERGVGWRLE--GKGGFPEP	VDVDCGNAGTGVRLIMGAAGFAMCATFTFGD	117
Qy	123	ASLTKRPMGRVNLPLRMGVQVKS	DGDRLPVTLRPGKFTPTITYRVPMA	SAQVKSAVLL 182

	Query Match	43.2%	Score 989.5;	DB 16;	Length 746;
	Best Local Similarity	48.5%;	pred. No. 6.3e-52;		
	Matches 214;	Conservative 62;	Mismatches 152;	Indels 13;	Gaps 3;
Qy	12 ARKSSGLSCTVRIICDKNISHSRSMFPGCLASGETIITGLLEGEDVINTCKMAQANGARIR	71	:	:	:
Dd	317 AQPGLSGTIRVPGDKNISHSRMGLSLAEGTTVEGLEGEDALATQAFDMGVVIE	376	:	:	:
Qy	72 -KEGDTWTIDGVNGNGLLAPLEPLDFGNAATCRLTMTGLVGVDYDFDFTIGDASLTKKRP	130	:	:	:

Db 377 GPQNGRVTVHGVGLHGLKAPPGPIYLLGNSGTSMLLSGLLAAQPFEDSTLTGDASLSKRP 436

Qy 131 GRVNLPLREMGVOVQKSEDDRLPVTLRGKPTPTITVTRVPMASQAQKSAVLLAGLNTPGI 190

Db 437 NRVAKPLREMGAVTETGEGRPMPPTIRGGQRLTGMDHMDPMSQAQKSAVLLAGLNTAA 496

Qy 191 TTVIEPTWTRDTEKMLQGGANLTVETDADGVRTILEGRGKLTGQVIDVPGDPSSTAF 250

Db 497 TSVTEPATRDRHTERMLTGRGFGVPVVEGS-----TAKVESGHKLSATHIEVPADISSAAF 551

Qy 251 PLVAALVPGSDVTILNVLNMPRTGLITLQEMGADIEVINPRLAGGEDVADLRVSST 310

Db 552 FLVAASTAEGSELVHQVGNPTRVGVIEILRLMGDLSLENQREVGEPVADIRVRSAR 611

Qy 311 LKGVTVPEDRAPSMIDEPYILAVAAAFAGATVNGLEELRVKESDRLSAVANGKLVNGV 370

Db 612 LKGDIDPELVLAIDEPFVLVFAAACAEGRVTLRGAEELRVKESDRISQVNMADGLKALGV 671

Qy 371 DCDEGETSLVVRGPRDGKGLGNASGAATVATHLDRHIAVMSFLVGLVSNPVTVDATMIA 430

Db 672 KAETPDGIVIEG-----GAFGGEVWAGHDHRIAMSFVSASLRASGPIRIHDCANVA 724

Qy 431 TSFPEFMDLMAGLGAKIELSD 451

Db 725 TSFNPFLALCAQTGIRAVEN 745

RESULT 6

Q9PB21 ID Q9PB21 PRELIMINARY; PRT; 454 AA.

AC Q9PB21

DT 01-OCT-2000 (Tremblrel. 15, Created)

DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)

DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)

DE 3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE.

GN XF2324.

OS Xylella fastidiosa.

OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;

OC Xylella.

OX NCBI_TaxID=2371;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=9A5C;

RX MEDLINE=20365717; PubMed=10910347;

RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,

RA Alvares R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,

RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,

RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carver H.,

RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,

RA Coutinho L.L., Cristofani A.J.S., Ferreira V.C.A., Ferro J.A.,

RA Facincani A.P., Ferreira A.J.S., Frohne M., Furlan L.R.,

RA Fraga J.S., Franca S.C., Franco M.C., Gomes S.L., Gruber A.,

RA Garner M., Goldman G.H., Goldman M.H.S., Kemp E.L., Kitajima J.P.,

RA Ho P.L., Hohnselt J.D., Junqueira M.L., Kasper E.L., Kitajima J.P.,

RA Krieger J.E., Kurame E.E., Laigret F., Lambais M.R., Leite L.C.C.,

RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,

RA Machado M.A., Madeira A.M.B.N., Martins E.M.F., Matsukuma A.Y.,

RA Marques M.V., Miracca E.C., Miyaki C.V., Monteiro-Vitorello C.B.,

RA Menck C.F.M., Miracca E.C., Miyaki C.V., Netto L.E.S.,

RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,

RA Nhani A.Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,

RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,

RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,

RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,

RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,

RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,

RA da Silva J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,

RA de Souza A.F., Terenzi M.F., Truffi D., Tsai S.M., Tsubako M.H.,

RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,

RA Zago M.A., zatz M., Meidanis J., Setubal J.C.,

RT "The genome sequence of the plant pathogen Xylella fastidiosa";

RL Nature 406:151-159(2000).

CC EMBL; AF004043; AAF85123.1; -

DR InterPro; IPR001986; EPSP_synthase.

DR Pfam; PF00275; EPSP_synthase; 1.

DR ProDom; PD001867; EPSP_synthase; 1.

DR PROSITE; PS00104; EPSP_SYNTHASE_1; UNKNOWN_1.

DR PROSITE; PS00885; EPSP_SYNTHASE_2; UNKNOWN_1.

KW Complete proteome.

SQ SEQUENCE 454 AA: 48266 MW: 45CCF074B6C0BA57 CRC64;

Query Match 39.9%; Score 912.5; DB 16; Length 454;

Best Local Similarity 46.9%; Pred. No. 1.5e-47;

Matches 206; Conservative 60; Mismatches 158; Indels 15; Gaps 4;

Qy 12 ARKSSGLSGTVRIIPGDKSISHRSMFEGGLASGETRITGLLEGEDVINTKAMQAMGARIR 71

Db 23 AHOGTPLHGVLSIPGDKSISHRVMAFAALADGTSRIDGFLAEADTCTABILARLGVRIE 82

Qy 72 KEGDTW-IDGVNGGLLAPAEPLDRGNATGCLRMGLVGVYDFDSTFTGDASLSKRP 130

Db 83 TPLSTQRIVHGVGVGDLQASHIPLDCGNAGTGMRLLAGLLVAQPFDSVLVGDASLSKRP 142

Qy 131 GRVNLPLREMGVOVQKSEDDRLPVTLRGKPTPTITVTRVPMASQAQKSAVLLAGLNTPGI 190

Db 143 RRTVDPLSQMGARIDTSDDGTPLRIYGGQLLHGIDFISPVASQAQKSAVLLAGLYARNE 202

Qy 191 TTVIEPTWTRDTEKMLQGGANLTVETDADGVRTILEGRGKLTGQVIDVPGDPSSTAF 250

Db 203 TVVREPHPTDRYTERMLTAFGVVIDVSTGC-----ARLRGGQRLCATDITIPADFSSAAF 257

Qy 251 PLVAALVPGSDVTILNVLNMPRTGLITLQEMGADIEVINPRLAGGEDVADLRVSST 310

Db 258 YLVAASVPGSDVTLRAVGLNPRRIGLLTVLRMGANIVESNRHEQGQEPVVDLRVRYAP 317

Qy 311 LKGVTVPEDRAPSMIDEPYILAVAAAFAGATVNGLEELRVKESDRLSAVANGKLVNGV 370

Db 318 LQTRVPEDLVADMIDEPFALVFAAAAEQTVVSGAAELRVKESDRLAAMVTGLRVLG 377

Qy 371 DCDEGETSLVVRGPRDGKGLGNASGAATVATHLDRHIAVMSFLVGLVSNPVTVDATMIA 430

Db 378 QVDETADGATTHGGPIGHGTINSHG-----DHRIAMAFSIAGOLSVSTVRIEDVANVA 430

Qy 431 TSFPEFMDL--MAGLGAKI 447

Db 431 TSFPEFMDL--MAGLGAKI 449

RESULT 7

Q9KCA6 ID Q9KCA6 PRELIMINARY; PRT; 431 AA.

AC Q9KCA6

DT 01-OCT-2000 (Tremblrel. 15, Created)

DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)

DT 01-OCT-2001 (Tremblrel. 18, Last annotation update)

DE 3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE (EC 2.5.1.19) (5-ENOLPYRUVYLHIKIMATE-3-PHOSPHATE SYNTHASE) (EPSP SYNTHASE) (EPSPS).

DE AROE OR BH1667.

GN Bacillus halodurans.

OS Bacteria; Firmicutes; Bacillus/Clostridium group;

OC Bacillus/Staphylococcus group; Bacillus.

OX NCBI_TaxID=86665;

RN [1]

RP SEQUENCE FROM N.A.

RX STRAIN=C-125 / JCM 9153;

RX MEDLINE=20512582; PubMed=11058132;

RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,

RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,

RA Horikoshi K.;

RT "Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and genomic sequence comparison with Bacillus subtilis";

RL Nucleic Acids Res. 28:4317-4331(2000).

CC -!- CATALYTIC ACTIVITY: PHOSPHOENOLPYRUVATE + 3-PHOSPHOSHIKIMATE -> ORTHOPHOSPHATE + O(5)-(1-CARBOXYVINYL)-3-PHOSPHOSHIKIMATE.

CC	-1- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS FROM CHORISMATE OF THE
CC	-1- AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC	-1- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
DR	EMBL; AP001512; BAB05386.1; -.
DR	InterPro; IPR001986; EPSP_synthase.
DR	Pfam; PF00275; EPSP_synthase; 1.
DR	ProDom; PD001867; EPSP_synthase; 1.
DR	PROSITE; PS00104; EPSP_SYNTHASE_1; 1.
DR	PROSITE; PS00885; EPSP_SYNTHASE_2; 1.
KW	Aromatic amino acid biosynthesis; Complete proteome; Transference.
SQ	SEQUENCE 431 AA; 45485 MW; 12E4FFBE7BA0743D CRC64;
	Query Match 38.6%; Score 884; DB 16; Length 431;
	Best Local Similarity 44.2%; Pred. No. 7.2e-46;
	Matches 188; Conservative 72; Mismatches 153; Indels 12; Gaps
QY	15 SSGLSGTVRIPGDKSIHRSFMFGGLASGETRITGLLEGEDVINTKAMQAMCARIKES 74 : : : : : : : : : : : :
QY	10 AKGLGTIKIVPGDKSIHRVMFEGALAKGTTTVEGPLDGADCLSTISCFQKLGVSTEQA 69 : : : : : : : : : : : :
Db	70 ERVTVKKGWDGLREPESDILDVNGSCTTTRLILGILSTLPFHVSIIIGDESICKRPMKRY 129 : : : : : : : : : : : :
QY	135 NPLREMGVQVKSED-GDRLPVTLRGPKPTPTTYRVPMASAQVKSAVLAGLNTPOITTV 193 : : : : : : : : : : : :
Db	130 EPLKSMGAQIDGRDHGNLTPLSRGGQL-KGIDFHPVASAQMKSAILLAGLAERCKTSV 188 : : : : : : : : : : : :
QY	194 IEPIMTDRHTKEMLQGFGANLVETDADGVTRTLRSGKGLTGQVTDVPGPSSTAFPLV 253 : : : : : : : : : : : :
Db	189 TEPAKTRDHTLERLEAFGNVI----EKDGL-TVSIEGQMLTGQHVVVPGDISSAFFLV 243 : : : : : : : : : : : :
QY	254 AALLVPGSDVTLNLVLNPNRTGLIILTQBMGDIEVINPRLAGGEDVADLRVRSTLK 313 : : : : : : : : : : : :
Db	244 AGAWPHSRITLTNVGINPTFRAGILEVLMQMATLAMENERVOGGEFPVADLLTIETSVLG 303 : : : : : : : : : : : :
QY	314 VTPVEDRAPSNIDEYPILAAVAFAFCATVWGNGLEELRVKESDRLSAVANGKLNGVDCD 373 : : : : : : : : : : : :
Db	304 VEIGGDIIPRLDEIPILIAVTAQASGRTVIKDAEELKVETNRIDTVVSELTCLKGASH 363 : : : : : : : : : : : :
QY	374 EGETSILVRCRDPCKGLGNASGAAVATHLDHRTAMSFYNGLVNSEPVTVDDATMIATSF 433 : : : : : : : : : : : :
Db	364 ATDDGMIIIEGTPFKG-----GYTVSSHGDRHRCMAIALLUAEKPVTVEGTEAIV 418 : : : : : : : : : : : :
QY	434 PEFMD 438 : : : : : : : : : : : :
QY	419 PSFSD 423 : : : : : : : : : : : :

```

RESULT      8
Q99283
ID      Q99283      PRELIMINARY;      PRT;      430 AA..
AC      C99283;
DT      01-JUN-2001 ('EMBLrel. 17, Created)
DT      01-JUN-2001 ('EMBLrel. 17, Last sequence update)
DT      01-DEC-2001 ('EMBLrel. 19, Last annotation update)
DE      PUTATIVE 3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE
DE      (EC 2.5.1.19).
OS      AROA OR SPY1352.
GN      Streptococcus pyogenes.
OC      Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC      Streptococcus.
OQ      NCBI_TaxID=1314;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=SF370 / ATCC 700294 / SEROTYPE M1;
RX      MEDLINE=21192684; PubMed=11296296;
RA      Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Santic G., Lyon K.,
RA      Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA      Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA      Yuan X., Clifton S.W., Roe B.A., McLaughlin R.
RT      *Complete genome sequence of an M1 strain of Streptococcus pyogenes.*

```

Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).	
EMBL; AE006573; AAK34180.1; -	
InterPro; IPR001986; EPSP_synstase.	
Pfam; PF002275; EPSP_synstase; 1.	
ProDom; PD001867; EPSP_synstase; 1.	
ProSITE; PS00104; EPSP_SYNTHASE_1; 1.	
ProSITE; PS00885; EPSP_SYNTHASE_2; UNKNOWN_1.	
KW Transferrase; Complete proteome.	
SEQUENCE 430 AA; 46692 MW; 327363B39020FB61 CRC64;	
Query Match	35.8%; Score 820; DB 16; Length 430;
Best Local Similarity	43.8%; Pred. NO. 5.4e-42;
Matches 189; Conservative 76; Mismatches 141; Indels 26; Gaps	
QY 13 RKSSG-LSGTVRIPGDKSISHRFSFGGLASGETRITGLLEGEDVINTGRAMAMGARIR 71	: : : :
DB 7 RTNAGPLOGTIOVPGDKASISHRAVILGAVAKGETRVKGLLGKEDVLSTIOAFRNIGVRIE 66	: : : :
QY 72 KEGDTWIIDGVNGGLLAPLBPALDFGNAATGCRLTWGLVGVYDFDSTFIGDASLTRKPMG 131	: : : :
DB 67 EKDBQVIEGGQFGLNAPQOTLNMNGSGTSMRLIAGLLAGQFVSVKMIGDESLSKRPMD 126	: : : :
QY 132 RVLNPLREMGVOYKSE--DGDRLPVTLRGKPTPTPIYRVPMAAQVKSALLAGLNTPGI 190	: : : :
DB 127 RIVYPLKMGVETISGETDRQFPPLQGGNRLNPITTYTLPISSAAQVKSAILLAALQAKGT 186	: : : :
QY 191 TTVTEPTMTDRHTEKMLQGFGANLTVTADGVRTIRLEGRGKLTGOVIDVPGDPSSTAF 250	: : : :
DB 187 TQVVEKEITRNHTEEMIQFGGRIV----DGRK-ITLVGPQQLTAQETIVPGDISNAF 241	: : : :
QY 251 PLVAALLVPGSDVITLVNLMNPTRTGLILTLQEMGADI--EVINPLRAGEDVADLRVRS 308	: : : :
DB 242 WLVAGLIIPGSELLKNVGNVPTRTGILEVVEKMGQAIVYEDNNKK-----EQVTSIRVY 297	: : : :
QY 309 STLKGVTVPEDRAPSMIDEPILAAVAAPFAGATVMNGLEELRVKESDRLSAVANGKLN 368	: : : :
DB 298 SNMKGTTISGGLIPRLIDELPITALTAQOGTCTCKDAQELRVKQETDRIQVVYTDILNSM 357	: : : :
QY 369 GVDCEGETSLVVRGRPDG---KGLGNASGAANAATHLDHRIAMSFVLMG--LVSENPVTVD 424	: : : :
DB 358 GAN-----IKATADGMIIKGPTVLYGANSTYGDHGRIGMTATAALLVKQGVHLD 408	: : : :
QY 425 DATMIATSPFEF 436	: : : :
DB 409 KEEAIMTSYPTF 420	: : : :
RESULT 9	
Q92A85 ID Q92A85 PRELIMINARY; PRT; 428 AA.	
AC Q92A85;	
DT 01-DEC-2001 (TREMBlrel. 19, Created)	
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)	
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)	
DE AROE PROTEIN.	
GN AROE OR LIN2037.	
OS Listeria innocua.	
OC Bacteria; Firmicutes; Bacillus/Clostridium group;	
OC Bacillus/Staphylococcus group; Listeria.	
OX NCBI_TaxID=1642;	
RN [1]	
RN SEQUENCE FROM N.A.	
RC STRAIN-CLIP 11262 / SEROVAR 6A;	
RX PubMed=11679669;	
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,	
RA Baquero F., Berche P., Bloecher H., Brandt P., Chakraborty T.,	
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,	
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,	
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,	
RA Gautier L., Goebel W., Gomez-Iopez N., Hain T., Hauf J., Jackson D.,	
RA Jones L.-M., Kaerst U., Krefit J., Kuhn M., Kunst F., Kurapkat G.,	
RA Madueno E., MailCournam A., Mata Vicente J., Ng E., Nedjari H.,	

```
RA Nordstiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001)
DR EMBL; AL5961170; CAC97287.1; -.
DR Listlist; LIN02037; -.
KW Complete proteome.
SQ SEQUENCE 428 AA; 45994 MW; 157B48C091A68FEB CRC64;

Query Match          35.6%; Score 815.5; DB 16; Length 428;
Best Local Similarity 40.8%; Pred. No. 1e-41;
Matches 173; Conservative 83; Mismatches 155; Indels 13; Gaps 5;

QY 17 GLSGTVRIPGDKSISHRSMFGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEDGT 76
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 9 GLVGETVPGDKSMHSRSMFCAIAEGKTVIRHFLRADDCLTGTFKAFKALGVKIEETDEE 68
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 77 WIIDGVNGGLLAPAPLDFGNAATGCRLTGMLGVGYDFDSTFFIGDASLTTRKPMGRVLP 136
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 69 IIVHGTGSDGLQAEGPLDIGNSGTITRLMGLGILAGRDFVTILGDESTAKRPMNRMLP 128
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 137 LREMGVQVKSDEGDRLLPVTLRGKPTPTTYRVPMAAQVSAVLLAGLNTPTGITTVE 195
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 129 LQEMGAKMKGKGSEFAPISIIQNSLKRMEYHMPVASAQVKSALFAALQAEGETIIE 188
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 196 PIMTRDHTKMLQGGFANLTVDADGVRTIRLEGRKLTGOVIDVPGDPSSTAFPLVNA 255
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 189 KEKTRDHTHEMRQFGG---ETEMDGL-TIRVKGQKFIGQEMTVPGDVSSAAFFIVAG 243
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 256 LLVPGSDVTLNVMNPTRTGLITLQEMGADIEVINPRLAGGEDVADLRVRSSTLKGYT 315
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 244 LIIPGSEIEIETHVGLNPTRTGIDFVVEQMGSGSLVYKDSRSTGKLAGTVVKSSELKGT 303
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 316 VPEDRAPSMIDEPILAVAAFAEGATVNMGLLELRVKESDRLSAVANGLKNGVDCDEG 375
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 304 IGGDIIPRLIDETPIVALLATQAGTTIIRDAELKVKETNRIDAVANELNKGADITPT 363
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 376 ETSLVVRGPRDGKGLGNASAAVATHLDRHIAFSLVMG-LVSENPVTVDATMTATSP 434
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 364 EDGLIIRGKTP-----LHAANVTSGDHRIGMQLQIAALLVEDGDVLDRAEAVSVSP 417
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 435 EPM 438
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 418 TFE 421
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 10
Q990Y6 PRELIMINARY; PRT; 428 AA.
AC Q990Y6;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE 3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE (EC 2.5.1.19).
GN AROA.
OS Enterococcus faecalis (Streptococcus faecalis).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Enterococcaceae;
OC Enterococcus.
OX NCBI_TaxID=1351;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OG19;
RA Huycke M.M., Shepard L., Joyce W., Wise P., Moore D.R., Gilmore M.S.;
RT "Forme fruste respiration by Enterococcus faecalis produces extracellular superoxide.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF318277; AAG53678.1; -.
DR InterPro; IPR001986; EPSP_synase.
DR Pfam; PF00275; EPSP_synase; 1.
DR ProDom; PD001867; EPSP_synase; 1.
DR PROSITE; PS00104; EPSP_SYNTHASE_1; UNKNOWN_1.
DR
```

```
DR PROSITE; PS00885; EPSP_SYNTHASE_2; UNKNOWN_1.
KW Transferase.
SQ SEQUENCE 428 AA; 45715 MW; 9E0F4FE4A893CA95 CRC64;

Query Match          34.2%; Score 781.5; DB 2; Length 428;
Best Local Similarity 40.5%; Pred. No. 1.1e-39;
Matches 173; Conservative 79; Mismatches 152; Indels 23; Gaps 7;

QY 18 LSGTVRIPGDKSISHRSMFGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEDGTW 77
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 12 LQGLTMLVPSDKSISHRSMFCAISSGKTTITNLRGEDCLTLAFLRSLGVNIEDDGTI 71
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 78 IIDGVNGGLLAPAPLDFGNAATGCRLTGMLGVGYDFDSTFFIGDASLTTRKPMGRVLP 137
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 72 TVEGRGFAGLKAKNTIDVNGSGTITRLMGLGILAGCPETRLAGDASIAKRPMMNRMLP 131
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 138 REMGVQVKS-EDGDRLLPVTLRGKPTPTTYRVPMAAQVSAVLLAGLNTPTGITTVEP 196
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 132 NQMGAEQGVQQTFFPISIRGTQNLQPIDTMPVASAQVKSALFAALQAEGTSVVVEK 191
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 197 IMTRDHTKMLQGGFANLTVDADGVRTIRLEGRKLTGOVIDVPGDPSSTAFPLVNAAL 256
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 192 EKTRDHTHEMRQFGGLEV---DG-KKIMLTGPOQLTGONVVVPGDISSAAFFELVAGL 246
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 257 LVPGSDVTLNVMNPTRTGLITLQEMGADIEVINPRLAGGEDVADLRVRSSTLK 312
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 247 VVPDSEILLKNVGLNQTRTGLIDVKNMGSGVTILN-----EDEANHSGLLVKTSOLT 300
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 313 GVTVPEDRAPSMIDEPILAVAAFAEGATVNMGLLELRVKESDRLSAVANGLKNGVDC 372
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 301 ATEIGGAIIPRLIDETPIVALLATQAGTTIIRDAELKVKETNRIDAVAKELTILGADI 360
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 373 DGETSLVVRGPRDGKGLGNASAAVATHLDRHIAFSLVMG-LVSENPVTVDATMTAT 431
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 361 PTDDGLIIRGPT-----SLHGRVTSYGDHRIGMQLQIAALLVKEGTVELDKAEVSV 414
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 432 SPPEFMD 438
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 415 SYPAFFD 421
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 11
Q990Y25 PRELIMINARY; PRT; 432 AA.
AC Q990Y25;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE 3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE.
GN AROA OR SAL297 OR SAV1464.
OS Staphylococcus aureus (Strain N315), and
OS Staphylococcus aureus (Strain Mu50).
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Staphylococcus.
OX NCBI_TaxID=158879, 158878;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.aureus (Strain N315), and S.aureus (Strain Mu50);
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-U I., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hiraoka H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus aureus";
RL Lancet 357:1225-1240(2001).
DR EMBL; AF003134; BAB42557.1; -.
DR EMBL; AF003362; BAB57626.1; -.
DR InterPro; IPR001986; EPSP_synase.
```

```
DR Pfam: PF00275; EPSP_synthase; 1.  
DR ProDom: PD001867; EPSP_synthase; 1.  
DR PROSITE: PS00104; EPSP_SYNTHASE_1; 1.  
DR TRANSFERASE; Complete proteome.  
KW TRANSFERASE; Complete proteome.  
SQ SEQUENCE 432 AA; 47068 MW; A42102057AD15C72 CRC64;  
  
Query Match 26.7%; Score 610; DB 16; Length 432;  
Best Local Similarity 34.0%; Pred. No. 2.8e-29;  
Matches 146; Conservative 80; Mismatches 181; Indels 22; Gaps 6;  
  
QY 15 SSGLSGTVRIPGDKSISHRSEFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEG 74  
DB 10 SGPLAGEIEVPCDKSWTHRAIMLASLAEGTSNIYPLLEGEDCRRTWDFRLGLGVDIKED 69  
  
QY 75 DTWIIDGVNGGLLAPEAPLDFGNAATGCRLTGMLGVYDFDSTFIGDASLTKRPMGRVL 134  
DB 70 DKLVVNSPGYKAFKTPHQVLYTGNSTGTRLLAGLLSGLGIESVLSDGVSIKGRPMGRVL 129  
  
QY 135 NPLREMGVQKSEGDRLPVTLRGPKTPTPIYRVPMASQVKSALLAGLNTPGITTVI 194  
DB 130 RPLKMDANIEGIDNYTPLIK-ESVINGINQMEVASAQVKSALLFASLFSNDTTVIK 188  
  
QY 195 EPIMRDHTKMLQGF----GANLTVDADGVRTIRLEGRGKLTGQVIDVPGDPSSTA 249  
DB 189 ELDSVRSNHTWFRNPPIEAERLSITTPDAIQHKPAD-----FHVPGDLSSAA 240  
  
QY 250 PFLVAALVPGSDVTILNVMNPTRLTLTQEMGADIEVINPLAGGEDVADLRVR-S 308  
DB 241 FFIVAALITPESDVTIHNVGINPTRSIGIIDIIVKMGNGNIOLFQ-QTGAEPASIRIQVT 299  
  
QY 309 STLKGVTPVEDRAPSWIDEPYILAAFAEGATVWNGLEELRVKESDKLSAVANGLKUN 368  
DB 300 PMLQPTITTEGELVPAIDELPVIALCTQAVGTSTIKDAEELKVKETNRIDTADMNL 359  
  
QY 369 GVDCEGETSLVVRGPRGDKGLGNAGAAVATHLDRHTAMSLFVNLGVSENPVTVDATM 428  
DB 360 GFELQPTNDGLIIHPSE-----FTNATVDSLTDRHIGMMLAVASLLSSEPVKIQFPA 413  
  
QY 429 IATSPPEFM 437  
DB 414 VNVSPPGFL 422  
  
RESULT 12  
Q9RHZ8 PRELIMINARY; PRT; 207 AA.  
DT 01-MAY-2000 (TremBLrel. 13, Created)  
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)  
DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)  
DE 5-ENOLPYRUVILSHIKMATE 3-P SYNTHASE (FRAGMENT).  
GN AROF.  
OS Pseudomonas stutzeri (Pseudomonas perfectomarina).  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.  
OX NCBI_TaxID=316;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-JM300;  
RX MEDLINE=92013931; PubMed=1919506;  
RA Fischer R.S., Zhao G., Jensen R.A.;  
RT "Cloning, sequencing, and expression of the P-protein gene (phea) of  
RT Pseudomonas stutzeri in Escherichia coli: implications for  
RT evolutionary relationships in phenylalanine biosynthesis.";  
RL J. Gen. Microbiol. 137:1293-1301(1991).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-JM300;  
RX MEDLINE=99298294; PubMed=10368439;  
RA Xie G., Bonner C.A., Jensen R.A.;  
RT "A probable mixed-function supraoperon in Pseudomonas exhibits gene
```

```
RT organization features of both intergenomic conservation and gene  
RT shuffling.";  
RL J. Mol. Evol. 49:108-121(1999).  
DR EMBL: AF038578; AAD47363.1; -.  
DR InterPro: IPR001986; EPSP_synthase.  
DR Pfam: PF00275; EPSP_synthase; 1.  
DR ProDom: PD001867; EPSP_synthase; 1.  
DR PROSITE: PS00104; EPSP_SYNTHASE_1; UNKNOWN_1.  
FT NON_TER 207 207  
SQ SEQUENCE 207 AA; 21780 MW; 7231191C72A21D6B CRC64;  
  
Query Match 20.6%; Score 470.5; DB 2; Length 207;  
Best Local Similarity 51.5%; Pred. No. 2.8e-21;  
Matches 103; Conservative 22; Mismatches 70; Indels 5; Gaps 2;  
  
QY 12 ARKSSGLSTVRIPGDKSISHRSEFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIR 71  
DB 10 ANPGSGLTQGLRVPGDKSISHRSIMLSLAEGTTEXEGFLEGEDALATXAQFRMGVVI- 68  
  
QY 72 KEG---DTWIIDGVNGGLLAPEAPLDFGNAATGCRLTGMLGVYDFDSTFIGDASLTKR 128  
DB 69 -EGPHQGRVTVHGVGLHGLQAPPPIYLGNSGTSMRLLAGLLAAQPFDTTLSGDASLTKR 127  
  
QY 129 PMGRVLNPLREMGVQKSEGDRLPVTLRGPKTPTPIYRVPMASQVKSALLAGLNT 188  
DB 128 PMNRVAKPLREMGAVTETAAGRPPITIRGGKKLSGMHYDMPMASQVKSCLLAGLYAA 187  
  
QY 189 GITTVEIETMRDHTKMLQ 208  
DB 188 GKTSVTEPAPTRDHTERMQQ 207  
  
RESULT 13  
Q9HQC1 PRELIMINARY; PRT; 439 AA.  
AC Q9HQC1;  
DT 01-MAR-2001 (TremBLrel. 16, Created)  
DT 01-MAR-2001 (TremBLrel. 16, Last sequence update)  
DT 01-JUN-2001 (TremBLrel. 17, Last annotation update)  
DE 3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE.  
PSC OR VNG1232G.  
OS Halobacterium sp. (strain NRC-1).  
OC Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae;  
OC Halobacterium.  
OX NCBI_TaxID=64091;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20504483; PubMed=11016950;  
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,  
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,  
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,  
RA Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,  
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,  
RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudis J.L., Jung K.-H.,  
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,  
RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;  
RT "Genome sequence of Halobacterium species NRC-1.";  
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).  
DR EMBL: AE005049; AAG19594.1; -.  
DR InterPro: IPR001986; EPSP_synthase.  
DR Pfam: PF00275; EPSP_synthase; 1.  
DR ProDom: PD001867; EPSP_synthase; 1.  
DR PROSITE: PS00104; EPSP_SYNTHASE_1; UNKNOWN_1.  
DR PROSITE: PS00885; EPSP_SYNTHASE_2; UNKNOWN_1.  
KW Transferase; Complete proteome.  
SQ SEQUENCE 439 AA; 44333 MW; 48CAD75A1F0CEB89 CRC64;  
  
Query Match 20.6%; Score 470.5; DB 17; Length 439;  
Best Local Similarity 31.3%; Pred. No. 7.8e-21;  
Matches 139; Conservative 76; Mismatches 190; Indels 39; Gaps 13;
```

Qy 16 SGLSTGVIPGDKSISHSRSMFGGLASGETRITGLLEGEDVINTGKAMQAMGARKEGD 75
Db 17 SRVRGRARAPSPKSYTHALLAAGYADGETVVRDPLVSADTRATARAVELLGGAAARENG 76
Qy 76 TWIDGVNGGGLLAPAPLDFGNAATGRLTMGLVGVYDFDSTFGDASLRKPMGRVLN 135
Db 77 DWVVTGFSRPAI-PDAVIDCANSQTTMRLVTAALADAGTTVLGTGDESRLRPHGPLLD 135
Qy 136 PLREMGVQVSKEDGD-RLPVTLRGPKTPTITYRVP-MASQVKSAYLLAGLNT-----P 188
Db 136 ALSGLGGTARTRNGQAPLVVDGVSQGSVA--LPGDVVSQFVALLMAGAVTGTET 193
Qy 189 GTTIVIEPIMTRDHEKMLQFGANLTVETDADGVRTIRLEGRG-----KLTCQVIDVPGD 244
Db 194 DLTTTELKSAPIVDITLDVDAFGVCA-ET-AAGYRV-----RGQAYAPSAEAVVPGD 246
Qy 245 PSSAFPLVALLVPGSDVTLINVLNMPNTRG---LIITLQEMGADIIEVNPRLAGGEDV 301
Db 247 FSSASYLLAAGALAAADGAAYVVEGMHPSAQGDAIIVDLVLERMGADID-----WDTES 299
Qy 302 ADLRVRSSTLKGVTVPEDRAPSMIDEXYPIAVAAAFAGATVMNGLEELRVKESDRLSAV 361
Db 300 GVITVQSELSEVGVADTDPOLL---PTIAVLGNAADGTTITDAEHVRYKETDRVAA 356
Qy 362 ANGLKLVGDCDEGETSLVVRGRDPKGLGNAGSAAVATHLDHRTAMSLVNGLYSENPV 421
Db 357 AESLSKLGASVEERDELVVRG-----GDTLSGASVDSGRGDRHLVMAVAGLVADGET 411
Qy 422 TVDDATMTATSPFPEMDLMAGLGA 445
Db 412 TIAGSEHVDVSPFPDFEVLGGLA 435

RESULT 14
Q9L213 PRELIMINARY; PRT; 440 AA.
ID Q9L213;
AC Q9L213;
DT 01-OCT-2000 (TremBLrel. 15, Created)
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TremBLrel. 17, Last annotation update)
DE 3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE.
GN ARCA.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA Brown S.P., Harris D.;
IL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kieser H.M., Denapaita D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmid and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL; AL138598; CAB71266.1;
DR InterPro; IPR001986; EPSP_synthase.
DR InterPro; IPR000408; RCCL.
DR Pfam; PF00275; EPSP_synthase.1.
DR ProDom; PD001867; EPSP_synthase.1.
DR PROSITE; PS00885; EPSP_SYNTHASE.2; 1.
DR PROSITE; PS00626; RCCL_2; UNKNOWN_1.
KW transferase.

SQ SEQUENCE 440 AA; 46529 MW; 77B720F81398EB0D CRC64;
Query Match 18.2%; Score 416.5; DB 2; Length 440;
Best Local Similarity 28.0%; Pred. No. 1.5e-17;
Matches 128; Conservative 78; Mismatches 198; Indels 53; Gaps 13;
Qy 11 TARKSSGLSTGVIPGDKSISHSRSMFGGLASGETRITGLLEGEDVINTGKAMQAMG--- 67
Db 15 TIRLTGDEIVRVLGSKSYNRYLAIASLQSEQVETIDNALLSDDTVVFSRAIETFGHVT 74
Qy 68 -----ARIKKEGTWIIDGVNGGGLLAPAPLDFGNAATGRLTMGLVGVYDFDSTFI 120
Db 75 CDIDHATARIR-----VTPTGR-PMRAPSEDIIVGGAGTPLRFLISMAGHADGTTTIT 126
Qy 121 GDASITKRPMPGRVLNPLREMGVQVSKEDGD-RLPVTLRGPKTPTITYRVPMSAQVKA 179
Db 127 GNARQERPMGDLKALPALGVADATAVRNGSSPPRVVVGSGFKGATSGISGAVSQFTSS 186
Qy 180 VLLAGLNPFGIT--TVIEPIMTRDHEKMLQFGANLTVETDADGVRTIRLEGRKLTGQ 237
Db 187 LIINALRAQTDTETITISDLSVKPYVEMTLA-AGLVSVDRDGYRRFTVPSGOQARG 245
Qy 238 VIDVPGDSSSTAFPLVAA-----LLVPGSDVTLINVLNMPNTRTGLTLTLOEMGADIEVI 291
Db 246 QVTVEPDASGMSYFLAAAILQSRVVIIPG-----IGAGSHQGDVHLVQALERMGRTVEV- 299
Qy 292 NPLRAGGEDVADLRVRSSTLKGVTVPEDRAPSMIDEXYPIAVAAAFAGATVMNGLEELR 351
Db 300 -----GDD--SITVGGPLRGIDIDMEAMPDV--FSLAIVAAVAEGTTRITNIASLR 348
Qy 352 VKESRLSAVANGKLVGDCDEGETSLVVRGRDPKGLGNAGSAAVATHLDHRTAMSL 411
Db 349 VKECDRIAAVTTELKMGIDVEHSDAMYITG-----GTPHGAVIDTYDDHRIAMTFA 401
Qy 412 VNGLYSENPVTVDDATMTATSPFPEMDLMAGLGA 448
Db 402 IGLRTEG-VVIKDPCGVAKSPAPFQWTLDTLHPDLE 437

RESULT 15
Q9KR80 PRELIMINARY; PRT; 426 AA.
ID Q9KR80
AC Q9KR80;
DT 01-OCT-2000 (TremBLrel. 15, Created)
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)
DE 3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE (EC 2.5.1.19) (5-
DE ENOLPYRUVYLSHIKIMATE-3-PHOSPHATE SYNTHASE) (EPSP SYNTHASE)
DE (EPSPS).
GN VC1732.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Unayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae";
RL Nature 406:477-483(2000).
CC -!- CATALYTIC ACTIVITY: PHOSPHOENOLPYRUVATE + 3-PHOSPHOSHIKIMATE -
CC ORTHOPHOSPHATE + O(5)-(1-CARBOXYVINYL)-3-PHOSPHOSHIKIMATE.
CC -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS FROM CHORISMATE OF THE
CC AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.

```
DR EMBL; AE004251; AAF94882.1; -.
DR TIGR; VC1732; -.
DR InterPro; IPR001986; EPSP_synthase.
DR Pfam; PF00275; EPSP_synthase; 1.
DR ProDom; PD001867; EPSP_synthase; 1.
DR ProSITE; PS00104; EPSP_SYNTHASE.1; 1.
DR ProSITE; PS00885; EPSP_SYNTHASE.2; 1.
KW Aromatic amino acid biosynthesis; Complete proteome; Transferase.
SQ SEQUENCE 426 AA; 46101 MW; 38852D6483BFE1C3 CRC64;

Query Match      17.5%; Score 401.5; DB 16; Length 426;
Best Local Similarity 28.0%; Pred. No. 1.1e-16;
Matches 125; Conservative 87; Mismatches 180; Indels 55; Gaps 15;

QY 18 LSGTVRIPGDKSISHSRSMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGD-- 75
Db 12 ISGEVNLPGSKSVSNRALLLAALASGTRTLNLLSDDIRHMLNALTCLGVNRYLSADKT 71

76 TWIIDGVNGGLLAPAPLDGNAATGCRITMGLVGVYDFDSTFIGDASLTKRPMGRVLN 135
Db 72 TCEVEGLGQAFHTTQPLEFLGNAGTAMRPLAAALCLGQGVYVLTGEPKMKERPIGHLVD 131

QY 136 PLREMGVQVKSDEGDRLP-----VTLRGPKTP---TPITYRVPMSAQVKS 179
Db 132 ALRQAGAQIEYLEQENFPRLRQCTGLQAGTVTIDGSISSQFLTAFILMSAPLAQGVTK 191

QY 180 VLLAGLNTFCITTVIEPIINTRDHTKMLQGFANLAVETDADGVRTIRLEGKGLT-GQV 238
Db 192 IVGELVSKPYI-----DITLHIMEQFGVQV--INHDIQEFVIPAGQSYVSPGQF 238

QY 239 IDVPGPSSTAPLVLAALLVPGSDVTILNVLNPTRTGLIL--TLQEMGADIEVINPRLA 296
Db 239 L-VEGDASSASY-FLAAAIKGEVKTGIGKNSIQGDIQFADALEKMGQAIE----- 289

QY 297 GGEDVADLRVRSTLKGTVVPEDRAPSMIDEXYPILAVAAAFAGATVMNGLEELRVKESD 356
Db 290 WGGDY--VIARRGELNAVLDLDFNHIP---DAAMTIATTALFAKGTTAIRNVYWRVKETD 344

QY 357 RLSSAVANGKLKNGVDCDEGETSLVYVRGPDGKGLGNASGAAVATHLDHRIAMSFYMGVLV 416
Db 345 RLAAAMATELRKVGATVEGEDFIVI--TPPTKLI----HAAIDYDDHRMAMCFSLVAL- 397

QY 417 SENPVTVDATMTATSPFEMDLMAGL 443
Db 398 SDTPVTINDPKCTSKTFPDYDKFAQL 424
```

Search completed: August 15, 2002, 14:05:43
Job time: 412 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
 Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 15, 2002, 13:57:41 ; Search time 56.25 Seconds
 (without alignments)
 898.464 Million cell updates/sec

Title: US-09-464-099A-3
 Perfect score: 2288
 Sequence: 1 MSHCASSRPATARKSSGLSG.....FMDLMAGLAKIELSDTKAA 455

Scoring table:
 BLOSUM62
 Gapop 10.0 , Gapext 0.5

searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : A_Geneseq_032802.*

1:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
2:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
3:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
4:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
5:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
6:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
7:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
8:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
9:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
10:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
11:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
12:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
13:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
14:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
15:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
16:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
17:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
18:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
19:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
20:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
21:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
22:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2288	100.0	455	13 AAR22300	Class II EPSPS enz
2	2288	100.0	455	18 AAW34683	Class II EPSP synt
3	2288	100.0	455	18 AAW24474	Class II EPSPS for
4	2288	100.0	455	19 AAW71609	Agrobacterium sp.
5	2288	100.0	455	22 AAE05053	Agrobacterium sp.
6	2282	99.7	455	13 AAR26449	CP4-EPSPS. Synthe
7	2282	99.7	455	19 AAR39426	CP4-EPSPS protein.
8	2282	99.7	527	22 AAW52214	EPSPS SEQ ID NO 3.
9	1906.5	83.3	449	13 AAR22301	Class II EPSPS enz
10	1906.5	83.3	449	13 AAR22302	Class II EPSPS enz
11	1906.5	83.3	449	18 AAW34684	Class II EPSP synt

12	1906.5	83.3	449	18 AAW34685	Class II EPSP synt
13	1906.5	83.3	449	18 AAW24479	Class II EPSPS use
14	1906.5	83.3	449	18 AAW24480	Class II EPSPS use
15	1906.5	83.3	449	19 AAW71611	Pseudomonas sp. st
16	1906.5	83.3	449	19 AAW71610	Achromobacter sp.
17	1906.5	83.3	449	22 AAE05054	Achromobacter sp.
18	1906.5	83.3	449	22 AAE05055	Pseudomonas sp. st
19	867.5	37.9	447	18 AAW34688	Class II EPSP synt
20	867.5	37.9	447	18 AAW24487	Class II EPSPS for
21	867.5	37.9	447	19 AAW71619	Synechocystis sp.
22	867.5	37.9	447	22 AAE05072	Synechocystis sp.
23	806	35.2	443	18 AAW34689	Class II EPSP synt
24	806	35.2	443	18 AAW24488	Class II EPSPS for
25	806	35.2	443	19 AAW71620	Dichelobacter nodo
26	806	35.2	443	22 AAE05073	Dichelobacter nodo
27	803.5	35.1	427	22 AAB48177	S. pneumoniae araA
28	803.5	35.1	427	22 AAB37094	Streptococcus pneu
29	803.5	35.1	427	22 AAB48500	Streptococcus pneu
30	798.5	34.9	427	20 AAW97388	5-enolpyruvylshiki
31	798.5	34.9	427	21 AAY56504	Streptococcus pneu
32	798.5	34.9	427	22 AAB31158	A 5-enolpyruvylsh
33	785	34.3	428	18 AAW34686	Class II EPSP synt
34	785	34.3	428	18 AAW24481	Class II EPSPS use
35	785	34.3	428	19 AAW71617	Bacillus subtilis
36	785	34.3	428	22 AAE05070	Bacillus subtilis
37	704.5	30.8	415	20 AAW97389	Protein encoded by
38	704.5	30.8	415	21 AAY56505	Streptococcus pneu
39	704.5	30.8	415	22 AAB31159	A 5-enolpyruvylsh
40	704.5	30.8	415	22 AAB48178	S. pneumoniae araA
41	704.5	30.8	415	22 AAB37095	Streptococcus pneu
42	704.5	30.8	415	22 AAB48501	Protein encoded by
43	577	25.2	430	18 AAW34687	Class II EPSP synt
44	577	25.2	430	18 AAW24482	Class II EPSPS use
45	577	25.2	430	19 AAW71618	Staphylococcus aur

ALIGNMENTS

RESULT 1

AAR22300

ID AAR22300 standard; Protein; 455 AA.

XX

AC AAR22300;

XX

DT 03-AUG-1992 (first entry)

XX

Class II EPSPS enzyme.

DE

XX

XX

KW Glyphosate tolerant bacteria; herbicides; glyphosate resistant plants;

KW 5'-enolpyruvylshikimate-3-phosphate synthase; weed control.

XX

OS Agrobacterium sp. strain CP4.

XX

XX

PN WO9204449-A.

XX

XX

PD 19-MAR-1992.

XX

PF 28-AUG-1991; 91WO-US06148.

XX

XX

PR 31-AUG-1990; 90US-0576537.

XX

XX

(MONS) MONSANTO CO.

PA Barry GF, Kishore GM, Padgett SR;

XX

XX

DR WPI; 1992-114356/14.

XX

XX

PT DNA encoding class II 5'-enol pyruvyl shikimate-3-phosphate

PT synthase - for producing plants and bacteria tolerant to

XX

XX

XX

PS glyphosate herbicides

PS Disclosure; Fig 3; 148pp; English.

[illegible]

QY	241	VPGDPSSTAPLVAALLVPGSVDTIINLVLMNPTRTGLIITLQEMGADIEVINPRLAGGED	300
DB	241	vpqdpstaifpvaallvpqsdvtilnvimnprrtglitlqemgadievlnprlagged	300
QY	301	VADLRVRSSTLKGVTVPEDRAPSMIDEXPTLAVAAFAEGATVMNGLLEELRVKESDRLSA	360
DB	301	vadlrvrssltkgvtpedrapsmideyptlavaaaafegatvmnglleeelrvkesdrisa	360
QY	361	VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRITAMSFVYMGVLVSENP	420
DB	361	vanglklngvdcdegetslvvrgrpdgkglgnasgaavathldhriamsflvmglvsenp	420
QY	421	VTVDATMIATSPPEPMDLMAGLAKIELSDTKAA	455
DB	421	vtvddatmiatspefmdlmaglgakielelsdtkaa	455
RESULT 3			
AAW	24474	AAW24474 standard; Protein; 455 AA.	
AAW	24474;		
XX	02-OCT-1997	(first entry)	
DE	Class II EPSPS for glyphosate resistant plant production.		
XX	5-enolpyruvylshikimate-3-phosphate synthase; EPSPS; Class II;		
KW	glyphosate resistant; transgenic plant; herbicide; shikimic acid;		
KW	fusion protein; soybean; corn; wheat; barley; cotton; sugarbeet;		
KW	tobacco.		
XX	Agrobacterium sp. strain CP4.		
XX	US56333435-A.		
XX	27-MAY-1997.		
XX	31-AUG-1990; 90US-0576537.		
XX	13-SEP-1994; 94US-0306063.		
PR	31-AUG-1990; 90US-0576537.		
PR	28-AUG-1991; 91US-0749611.		
XX	(MONS) MONSANTO CO.		
PA	Barry GF, Kishore GM, Padgett SR, Stallings WC;		
PI	WPI: 1997-297418/27.		
XX	N-PSDB; AAW77313.		
XX	New isolated 5-enol:pyruvyl:shikimate-3-phosphate synthase gene -		
PT	used for transforming plants to produce plants which are tolerant to		
PT	glyphosate herbicide		
XX	Claim 1; Column 57-60; 154pp; English.		
PS	AAW24474 shows the sequence of a class II 5-enolpyruvylshikimate-3-		
CC	phosphate synthase (EPSPS) enzyme. Class II EPSPS enzymes are tolerant		
CC	to glyphosate herbicides. EPSPS and sequences encoding it are used for		
CC	the production of herbicide resistant (glyphosate-tolerant) plants		
CC	such as corn, wheat, rice, barley, soybean, cotton, sugarbeet, oilseed		
CC	rape, canola, flax, sunflower, potato, tobacco, tomato, alfalfa, poplar		
CC	pine, eucalyptus, apple, lettuce, peas, lentils, grape and turf grass		
XX	Sequence 455 AA;		
SQ			
Query Match 100.0%; Score 2288; DB 18; Length 455;			
Best Local Similarity 100.0%; Pred. No. 6.2e-185;			
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps			
QY	1	MSHGASSRPATARKSSGLSGTVPIDGDKTSHRSFMFGGLASGETRITGLLEGEDVINTG	60

[illegible]

Query Match	100.0%;	Score 2288;	DB 18;	Length 455;	XX
Best Local Similarity	100.0%;	Pred. No. 6.2e-185;			PT
Matches 455;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	PT
					XX
QY	1	MSHGASRPATARKSSGLSGTVRIPGDKSIHSRSMFEGGLASGETRITGTLLEGEDVINTG	60		PS
					XX

Glyphosate resistant 5-enol:pyruvyl:shikimate-3-phosphate synthase
 useful for characterisation of the enzyme to determine inhibition
 data values
 Claim 3; Fig 3; 152pp; English.

Db 421 vtvdattmatsfpefmdlmaglakielstkaa 455

RESULT 6

AAAR26449
ID AAR26449 standard; Protein; 455 AA.

XX AAR26449;

AC AAR26449;

DT 28-JAN-1993 (first entry)

XX CP4-EPSPS.

DE PMONI1030; CTP2; CP4; EPSPS; chloroplast transit peptide;

XX 5-enolpyruvyl-3-shikimate phosphate synthetase; ACC; ethylene;

KW Pseudomonas chloroaphis 6G5.

XX Synthetic.

OS WO9212249-A.

XX 23-JUL-1992.

XX 17-DEC-1991; 91WO-0509437.

XX 26-DEC-1990; 90US-0632440.

XX (MONS) MONSANTO CO.

XX Kishore GM, Klee HJ;

PI WPI; 1992-284334/34.

DR N-PSDB; AAQ27201.

XX Delaying fruit ripening and senescence in plants - by controlling

PT ethylene prodn., pref. by expression of

PT 1-amino:cyclopropane-1-carboxylic acid deaminase

XX Disclosure; Page 64-66; 110pp; English.

PS The sequences given in AAR26448 and AAR26449 are encoded by genes which

CC were used in the construction of an expression plasmid pMON11030 which

CC was used to transform petunia plants. This plasmid contained the genes

CC for chloroplast transit peptide (Ctp2) (AAR26448) and the CP4 synthetic

CC 5-enolpyruvyl-3-shikimate phosphate synthetase (EPSPS) gene (AAR26449)

CC which is capable of conferring resistance to glyphosate. The plasmid

CC also contained the 1-aminocyclopropane-1-carboxylic acid (ACC)

CC deaminase gene from Pseudomonas chloroaphis (see also AAQ27199). The

CC transformed plants had ethylene levels reduced to about one half that

CC of the control, untransformed plants. It is expected that such plants

CC will show reduced senescence of flowers and leaves when compared to

CC untransformed plants.

XX Sequence 455 AA;

PS Query Match 99.7%; Score 2282; DB 13; Length 455;

CC Best Local Similarity 99.8%; Pred. No. 2e-184;

CC Matches 454; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSHGASSRPTARKSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60

Db 1 mlhgassrptarksglsgrvripgdkshrsfmfgglasgetritgllegedvintg 60

QY 61 KAMQAMGARIRKEGDTWIIDGVNGGLLAPAPLDGFGNAATGCRLLTMGLVGVYDFDSTFI 120

Db 61 kamqamgarirkegtwiidgvnggllapeapldfngaatgcrlltmglvgvydfdstfi 120

QY 121 GDASLTKRPMGRVNLNPLREMGVQVKSSEDGRLPVTILRGKTPPTPIYRVPMASQVKSVA 180

Db 121 gdasltkrpmgrvnlplrengvqvksgdgrlpvtlrgpktptityrvpmasqvkstav 180

QY 181 LLAGLNTPGITTTIEPIIMTRDHTKMLQGGFANLTVETDADGVRTIRLEGRGLTGQVID 240

Db 181 lllaglnptgittviepimtrdhtekmlqggfanltvetdadvrtirlegrgkltgqvld 240

QY 241 VPGDPSSTAFPLVAALLVPGSDVYTLNVLNMPRTGTLILTLQEMGADIEVINPRLAGGED 300

Db 241 vpgdpsstafplvaallvpgsdvtilnvlmptrtgtiltlqemgadielvinprrlagged 300

QY 301 VADLRVRSSTLKGTVTPEDRAPSMIDEPVTLAVAAAFAGATVMNGLEELRVKESDRLSA 360

Db 301 vadlrvsrstlkgvtvpedrapsmidvptilavaaaafagatvmngleelrvkesdrisa 360

QY 361 VANGKLNGVDCDEGETSLVVRGRPDGKGLGNASGAATVATHLDRHRIAMSFVLMGLVSEN 420

Db 361 vangklngvdcdegetslvvrgrpdgkglgnasgaatvathldhriamsvflmglvsenp 420

QY 421 VTVDATMIATSPFDFMDLMAGLGAKEIELSDTKAA 455

Db 421 vtvdattmatsfpefmdlmaglakielstkaa 455

RESULT 7

AAW39426

ID AAW39426 standard; Protein; 455 AA.

XX AAW39426;

AC AAW39426;

DT 19-MAY-1998 (first entry)

XX CP4-EPSPS protein.

DE 1-aminocyclopropane-1-carboxylic acid deaminase; ACC deaminase;

XX fruit ripening; ethylene production; glyphosate resistance;

KW 5-enolpyruvyl-3-shikimate phosphate synthase; EPSPS; CP4.

XX Synthetic.

OS US5702933-A.

XX 30-DEC-1997.

XX 06-NOV-1995; 95US-0553943.

XX 17-DEC-1991; 91US-0809457.

XX 26-DEC-1990; 90US-0632440.

XX 06-NOV-1995; 95US-0553943.

XX (MONS) MONSANTO CO.

XX Kishore GM, Klee HJ;

PI WPI; 1998-076419/07.

DR N-PSDB; AAV09719.

XX Production of plants with delayed ripening - using DNA encoding

XX 1-amino:cyclopropane-1-carboxylic acid deaminase

XX Example 9; Fig 21; 56pp; English.

PS This sequence represents the synthetic CP4 5-enolpyruvyl-3-shikimate

CC phosphate synthase (EPSPS) gene which is capable of conferring

CC resistance to glyphosate and is used in a novel method for producing

CC fruit-bearing plants with delayed ripening. The method involves the

CC expression of a 1-aminocyclopropane-1-carboxylic acid (ACC) deaminase

CC gene in a plant at a level sufficient to reduce ethylene production in

CC the fruit.

XX Sequence 455 AA;

QY Query Match 99.7%; Score 2282; DB 19; Length 455;

CC Best Local Similarity 99.8%; Pred. No. 2e-184;

CC Matches 454; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSHGASSRPATARKSSGLSGTVIRIPGDKSISHRSFPMFGLASGETRITGLLEGEDVINTG 60
Db 1 mlhgassrpataarkssglsgrtvpigdkssishrsfpmfglasgetritgllegedvintg 60
QY 61 KAMQAMGARIRKEGDTWIIDGVNGGLLAPEAPLDFGNAATGCRLTMTGLVGYVDFDSTFI 120
Db 61 kamqamgarirkegdtwiidgvnggllapeapldfgnaatgcrltmtglvgyvdfdstfi 120
QY 121 GDASLTKRPMGRVNLPLREMGVQVKSDEGDRPLVTLRGPKTPTTYRVPMSAQVKSAY 180
Db 121 gdasltkrpmgrvnlplremgvqvksedgdrplvtlrgpktpityrvpmsaqvksav 180
QY 181 LLAGLNTPGITTVIEPIIMTRDHTKMLQGFANLTVETDADGVRTIRLGRGKLTGOVID 240
Db 181 llaglntpgittviepimtrdhtkmlqgfannltvetdadvrtirlegrgkltgqvvid 240
QY 241 VPGDPSSTAFPLVAALLVPGSDVTILNVLNPTRTGLIILTQEMGADIEVINPRLAGGED 300
Db 241 vpgdpsstaflvaallvpgsdvtlinvlnpstrtgliltqemgadievlnprlagged 300
QY 301 VADLRVRSSTLKGTVTPEDRAPSMIDEYPIILAVAAFAEGATVMNGLEELRVKESDRLSA 360
Db 301 vadlrvrssstlkgvtvpedrapsmideypilavaafaegatvmngleelrvkesdrlsa 360
QY 361 VANGKLINGVDCDEGETSLVVRGPRDGKGLGNASGAATAVATHLDHRIAMSFVLMGLVSENP 420
Db 361 vangklngvdcdegetsllvvrgrpdgkglgnasgaatavathldhriamsflvmglvsenp 420
QY 421 VTVDATMIATSPPEPMDLMAGLAKIELSDTKAA 455
Db 421 vtvdattmiatspfePMDlmaglakielstdtkaa 455

RESULT 8
AAM52214
ID AAM52214 standard; Protein; 527 AA.
XX AAM52214;
XX AAM52214;
DT 12-FEB-2002 (first entry)
XX EPSPS SEQ ID NO 3.
XX Transgenic plant; herbicide resistance; EPSPS; GOX; Petunia hybrida;
KW 5-enolpyruvylshikimate-3-phosphate synthase; glyphosate oxidoreductase;
KW protoporphyrinogen IX oxidase; immunoglobulin; Agrobacterium; strain CP4;
KW chloroplast transit peptide.
XX Petunia hybrida.
OS Agrobacterium sp.
PN JP2001190168-A.
XX 17-JUL-2001.
XX 27-OCT-2000; 2000JP-0328811.
XX 29-OCT-1999; 99JP-0310244.
XX (SUMO) SUMITOMO CHEM CO LTD.
XX WPI; 2001-605307/69.
DR N-PSDB; ABA02854.
XX New herbicide-resistant plant -
XX Examples; Page 46-48; 66pp; Japanese.
XX The invention relates to a transgenic plant which shows resistance to a
CC herbicide in an amount inhibiting natural
CC 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) activity of the plant
CC and having at least one enzymatic activity selected from:
CC (1) EPSPS activity different from natural EPSPS activity of the plant or

CC (2) glyphosate oxidoreductase (GOX) activity different from the natural
CC GOX activity of the plant; and in which a gene encoding a protein having
CC the following properties: (a) combines specifically with a substance
CC participating to the herbicidal activity a herbicide of
CC protoporphyrinogen IX oxidase inhibiting type; (b) has substantially no
CC denaturing activity on a substance to which said protein combines
CC specifically; and (c) contains substantially no framework region of the
CC variable region of immunoglobulin. The present sequence is that of the
CC Petunia hybrida EPSPS chloroplast transit peptide and the Agrobacterium
XX sp. strain CP4 EPSPS gene.
SQ Sequence 527 AA;
Query Match 99.7%; Score 2282; DB 22; Length 527;
Best Local Similarity 99.8%; Pred No. 2, 5e-184;
Matches 454; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MSHGASSRPATARKSSGLSGTVIRIPGDKSISHRSFPMFGLASGETRITGLLEGEDVINTG 60
Db 73 mlhgassrpataarkssglsgrtvpigdkssishrsfpmfglasgetritgllegedvintg 132
QY 61 KAMQAMGARIRKEGDTWIIDGVNGGLLAPEAPLDFGNAATGCRLTMTGLVGYVDFDSTFI 120
Db 133 kamqamgarirkegdtwiidgvnggllapeapldfgnaatgcrltmtglvgyvdfdstfi 192
QY 121 GDASLTKRPMGRVNLPLREMGVQVKSDEGDRPLVTLRGPKTPTTYRVPMSAQVKSAY 180
Db 193 gdasltkrpmgrvnlplremgvqvksedgdrplvtlrgpktpityrvpmsaqvksav 252
QY 181 LLAGLNTPGITTVIEPIIMTRDHTKMLQGFANLTVETDADGVRTIRLGRGKLTGOVID 240
Db 253 llaglntpgittviepimtrdhtkmlqgfannltvetdadvrtirlegrgkltgqvvid 312
QY 241 VPGDPSSTAFPLVAALLVPGSDVTILNVLNPTRTGLIILTQEMGADIEVINPRLAGGED 300
Db 313 vpgdpsstaflvaallvpgsdvtlinvlnpstrtgliltqemgadievlnprlagged 372
QY 301 VADLRVRSSTLKGTVTPEDRAPSMIDEYPIILAVAAFAEGATVMNGLEELRVKESDRLSA 360
Db 373 vadlrvrssstlkgvtvpedrapsmideypilavaafaegatvmngleelrvkesdrlsa 432
QY 361 VANGKLINGVDCDEGETSLVVRGPRDGKGLGNASGAATAVATHLDHRIAMSFVLMGLVSENP 420
Db 433 vangklngvdcdegetsllvvrgrpdgkglgnasgaatavathldhriamsflvmglvsenp 492
QY 421 VTVDATMIATSPPEPMDLMAGLAKIELSDTKAA 455
Db 493 vtvdattmiatspfePMDlmaglakielstdtkaa 527

RESULT 9
AAR22301
ID AAR22301 standard; Protein; 449 AA.
XX AAR22301;
AC AAR22301;
DT 03-AUG-1992 (first entry)
XX Class II EPSPS enzyme.
DE Class II EPSPS enzyme.
XX Glyphosate tolerant bacteria; herbicides; glyphosate resistant plants;
KW 5'-enolpyruvylshikimate-3-phosphate synthase; weed control.
XX Achromobacter sp. strain LBAA.
XX WO9204449-A.
XX 19-MAR-1992.
XX 28-AUG-1991; 91WO-US06148.
XX 31-AUG-1990; 90US-0576537.

```

XX (MONS ) MONGANTO CO.
XX
XX Barry GF, Kishore GM, Padgett SR;
XX
XX WPI; 1992-114356/14.
XX
XX DNA encoding class II 5'-enol pyruvyl shikimate-3-phosphate
XX synthase - for producing plants and bacteria tolerant to
XX glyphosate herbicides
XX
XX Disclosure; Fig 5; 148pp; English.
XX
XX The sequence is that of the Class II 5'-enolpyruvylshikimate-3
XX phosphate synthase enzyme (EPSP) of Achromobacter sp. strain LBAA
XX CC It is used to create glyphosate resistant plants or seeds which
XX CC can be planted in a field of crops to selectively control weeds.
XX CC The crops selected for are e.g. corn, wheat, rice, oilseed rape,
XX CC tobacco and alfalfa. This provides a cost effective, environmentally
XX CC compatible weed control device. See also AAR22300 and AAR22302.
XX
XX SQ Sequence 449 AA;

Query Match 83.3%; Score 1906.5; DB 13; Length 449;
Best Local Similarity 83.1%; Pred. No. 1.le-152;
Matches 374; Conservative 32; Mismatches 41; Indels 3; Gaps

QY 1 MSHGASSRPATARKSSGLSGTVIRPGDKSIHSRSMFGGLASGETRITGLLEGEDVINTG 60
DB 1 mshsaspkptarrsealtgeiripgdksisrsmfmgglasgetritgllegedvintg 60
QY 61 KAMQAMGARIRKEGDWTIDGVNGGLLAPEDFGNAATGCRLTMLGLGVYDFDSTFI 120
DB 1 ramqamgakirkegdvwiingvngcllqpeaaldfgnagtgarltmglvgtymkcsfi 120
QY 61 ramqamgakirkegdvwiingvngcllqpeaaldfgnagtgarltmglvgtymkcsfi 120
DB 1 ramqamgakirkegdvwiingvngcllqpeaaldfgnagtgarltmglvgtymkcsfi 120
QY 121 GDASLTRKPRMGRVNLPIREMGVQVQKSEGDGRPLVTLRGPKTPPTITYRVPMAAQVKSAY 180
DB 121 gdaslkrpmgrvlnpiremgvqvqeaadgrmpitllgpktpanpityrvpmasqvksav 180
QY 181 LLAGLNTPGITVTLEPTWRDHTKMLQGFGANLTVETDADGVFTIRLEGCKLTGQVID 240
DB 181 llaglnpcgvttvleptwrdrhtekmlqggfagdlitvtdkdgvzhrirtgqgklvggtid 240
QY 241 VPGDPSSTAFPLVAALLVPGSDVTILNVLNMPRTGILTLQEMGADIEVINPRLAGGED 300
DB 241 vpgdpsstafplvaallvegsvdtrirnlmnprrtgitllqemgadielvnarlaged 300
QY 301 VADLRVRSSTLKGVTVPEDRAPSMIDEPYILAVAAAFAGATVNMGLEELRVKESDRLSA 360
DB 301 vadlrvaslklgvvpperapsmiddeypvllaasfaegtvmtdgldelrvksdrlla 360
QY 361 VANGKLINGVDCDEGTVSLVVRPGDCKGLGNASGAAVATHLDRIAMSLFVGLVSEN 417
DB 361 vangleangvdcdegtslvvrpgdckglgnasgaavathldhriamsflvmglvse 417
QY 421 VTVDATMIATSPFEFMDLMAGLAKIELS 450
DB 418 vtvdsmiatspfefmdmmpglgakiels 447

RESULT 10
AAR22302
ID AAR22302 standard; Protein; 449 AA.
XX
XX AC AAR22302;
XX
XX 03-AUG-1992 (first entry)
XX
XX DE Class II EPSPS enzyme.
XX
XX KW Glyphosate tolerant bacteria; herbicides; glyphosate resistant plants
XX
XX KW 5'-enolpyruvylshikimate-3-phosphate synthase; weed control.

```

XX	Pseudomonas sp. strain PG2982.
XX	WO9204449-A.
PN	
XX	19-MAR-1992.
PD	
XX	
XX	28-AUG-1991; 91WO-US06148.
XX	
PR	31-AUG-1990; 90US-0576537.
XX	
XX	(MONS) MONSANTO CO.
PA	
XX	Barry GF, Kishore GM, Padgett SR;
PI	WPI; 1992-114356/14.
XX	
DR	
XX	DNA encoding class II 5'-enol pyruvyl shikimate-3-phosphate
PT	synthase - for producing plants and bacteria tolerant to
PT	glyphosate herbicides
XX	
PS	Disclosure; Fig 7; 148pp; English.
XX	
CC	The sequence is that of the Class II 5'-enolpyruvylshikimate-3
CC	phosphate synthase enzyme (EPSP) of Pseudomonas sp. strain PG2982
CC	It is used to create glyphosate resistant plants or seeds which
CC	can be planted in a field of crops to selectively control weeds.
CC	The crops selected for are e.g. corn, wheat, rice, oilseed rape,
CC	tobacco and alfalfa. This provides a cost effective, environmentally
CC	compatible weed control device. See also AAR22300 and AAR22301.
XX	
SQ	Sequence 449 AA;
Query Match 83.3%; Score 1906.5; DB 13; Length 449;	
Best Local Similarity 83.1%; Pred. No. 1.1e-152;	
Matches 374; Conservative 32; Mismatches 41; Indels 3; Gaps	
QY	1 MSHGASSRPATARKSSGLSGTVRTPGDKTSIHSRFMGGLASGETRTITGLGEDVINTG 60
DB	1 mshsasbpataarseatlgeiripgdkshrsfmfgglasgetritlglegedvintg 60
QY	61 KAMQAMGARIRKEGDWTIIDGVNGGLLAPPEALDFGNAAATCRLTMGLVGVDFOSTFI 120
DB	61 ramqamgakirkegdvwlingvgngcillqpeaaldfnagtgarltnmglvydmktsfi 120
QY	121 GDASLTRPMGRVLNPLIREMGVQVKSEDDRELPLYLRGPKTPTPTTYRVPMASAQKSAV 180
DB	121 gdaslskrpmgrvinpliremgvgveaadgdrmpitligpktanpitrvpmasaqksav 180
QY	181 LLAGLNTPGITVTIEPTWRDHTKMLQGFGANLTVFETADGVRTIRLEGRGLTKGOVID 240
DB	181 llaqlntpgvtvteptwrtldhtekmlqqfgyadltvetdkgvrhritggklvgqtid 240
QY	241 VPGPDSSTAFLVAALLVPGSDDVTILNVLANPRTTGLILTLOEMGADIEVINPLAGGED 300
DB	241 vpgpdstaflvaallvegsvdtirnvlmpnptrtgitlltlqemgadievlnarlagged 300
QY	301 VADLRVRSSTLKGVTVPEDRAPSMIDEPYLAVAAAFAEGATVVMGLEELRVKESDRLSA 360
DB	301 vadlrvrasklkgvvpverapasmideyvpvlaaasfaegetvmddgldelrvkesdrflaa 360
QY	361 VANGKLKNGVDCDEGEVSLVVGRPCDGKGLGNASGAATAVTHLDHRIAMSFLVMCLVSEN 420
DB	361 vargleangvdcdegeamlsvrgrpdgkgig ---ggvtathldhriamsflvmglaaekp 417
QY	421 VTVDADTAIATSFPEFMDLAGLKAKIELS 450
DB	418 vtvdsmiatstsfefndmmpglgakielS 447
RESULT 11	
AAW34684	

RESULT 11
AAW34684

ID XX AAW34684 standard; Protein; 449 AA.
AC AAW34684;
DT 17-FEB-1998 (first entry)
DE Class II EPSP synthase (EPSPs) from *Achromobacter* sp. strain LBAA.
XX 5-enolpyruvylshikimate-3-phosphate synthase; EPSPs; glyphosate;
KW shikimic acid pathway; phosphoenolpyruvic acid; 3-phosphoshikimic acid;
KW 5-enolpyruvyl-3-phosphoshikimic acid; herbicide; transgenic plant;
KW glyphosate resistance gene; glyphosate-tolerance; promoter.
XX
OS *Achromobacter* sp. strain LBAA.
FH Key Location/Qualifiers
FT Region 200..204
FT /label= characteristc_region
FT /note= "see AAW34690"
FT Region 26..29
FT /label= characteristc_region
FT /note= "see AAW34691"
FT Region 173..177
FT /label= characteristc_region
FT /note= "see AAW34692"
FT Region 271..274
FT /label= characteristc_region
FT /note= "see AAW34693"
XX
PN US5627061-A.
XX 06-MAY-1997.
XX 07-JUN-1995; 95US-0476008.
XX 13-SEP-1994; 94US-0306063.
PR 31-AUG-1990; 90US-0576537.
PR 28-AUG-1991; 91US-0749611.
XX
PA (MONS) MONSANTO CO.
XX
PI Barry GF, Kishore GM, Padgett SR, Stallings WC;
XX WPI; 1997-271315/24.
DR N-PSDB; AAT93789.
XX
PT Production of glyphosate-herbicide tolerant plants - using DNA
PT encoding class II 5-enol:pyruvyl:shikimate-3-phosphate synthase
PT enzyme(s)
XX
DS Disclosure; Fig 4; 151pp; English.
XX
CC AAW34683-89 represent a new class of glyphosate-tolerant
CC 5-enolpyruvylshikimate-3-phosphate synthases (EPSPs). These novel
CC EPSPs enzymes have little homology with known Class I EPSPs enzymes, and
CC belong to a new class, Class II. The present sequence was isolated from
CC *Achromobacter* sp. strain LBAA. The EPSPs enzymes are part of the
CC shikimic acid pathway, which leads to the biosynthesis of aromatic
CC compounds. EPSPs converts phosphoenolpyruvic acid (PEP) and
CC 3-phosphoshikimic acid to 5-enolpyruvyl-3-phosphoshikimic acid, and is
CC inhibited by the herbicide glyphosate. It would be useful to produce
CC transgenic crops containing glyphosate resistance genes so that
CC glyphosate-containing herbicides can be applied to selectively kill
CC weeds. The novel EPSPs enzymes exhibit a low Km for PEP and a high Ki for
CC glyphosate, such that when introduced into a plant, the plant is made
CC glyphosate-tolerant, and EPSPs enzyme activity is not affected. These
CC Class II EPSPs enzymes are fused to a chloroplast transit peptide to
CC target the protein into the chloroplast, which is the site for the
CC shikimic acid pathway. In addition, the EPSPs gene is cloned into a plant
CC under the control of a promoter such as figwort mosaic virus promoter or
CC the cauliflower mosaic virus promoter, so that expression is enhanced.
XX
SQ Sequence "449 AA;
".

Query Match 83.3%; Score 1906.5; DB 18; Length 449;
Best Local Similarity 83.1%; Pred. No. 1.le-152;
Matches 374; Conservative 32; Mismatches 41; Indels 3; Gaps 1;
QY 1 MSHGASRPATARKSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
DB 1 mshaspkpatarrsealtgeiripgdkshrsfmfgglasgetritgllegedvintg 60
QY 61 KAMQAMGARIRKEDGTWIIDGVNGGLLAPAPLDGFNAATGCRLTMGLVGVYDFDSTFI 120
DB 61 ramdamgakirkedgvwllngvngccllqpeaaldfnagtgartlmglyvtyamktsf 120
QY 121 GDASLTRPMGRVLNPLREMGVQVKSEGDRLPVTLRGPKTPPTPIYRVPMASAOVKSAY 180
DB 121 gdaslskrpmgrvlnplremgvqvveaagdrmpitllgpktnpityrvpmasaqvksay 180
QY 181 LLAGINTPGITTVIEPIWTRDHTKMLQGFANLTVEVDAGVTRIRLEGRKLTGOVID 240
DB 181 lllagintpgvttviepvmtrdhtekmlqgfgaditvetdkdvrrhrritgggkllvgqcid 240
QY 241 VPGDPSSTAFPLVAALLVPGSDVTTLNVLNMPTRTGLTLTQEMGADIEVINPRLAGGED 300
DB 241 vpgdpsstafplvaallvegsvdvtirvlnmptrtgltlitqemgadielvinarlagged 300
QY 301 VADLRVRSSTLKGTVVPEDRAPSMIDEYPIILAVAAFAEGATVMNGLEELRVKESDRLSA 360
DB 301 vadlrvrasklkgvvvpprapsmideypvllaiaasfaegetvmdgldeirvkesdr 360
QY 361 VANGKLNGVDCDEGETSLIVYRGDPDGKGLGNASAAVATHLDHRIAMSFVLMGLVSENP 420
DB 361 vargleangvdcgetsemsltvrgprdgkglg---ggvtathldhriamsflvmglaaekp 417
QY 421 VTVDATMIATSFPEFMDLMAGLGAKIELS 450
DB 418 vtvdsmniatsfpefmdmmpglgakiels 447
RESULT 12
AAW34685
ID AAW34685 standard; Protein; 449 AA.
XX
AC AAW34685;
XX
DT 17-FEB-1998 (first entry)
XX Class II EPSP synthase (EPSPs) from *Pseudomonas* sp. strain PG2982.
DE
DE 5-enolpyruvylshikimate-3-phosphate synthase; EPSPs; glyphosate;
KW shikimic acid pathway; phosphoenolpyruvic acid; 3-phosphoshikimic acid;
KW 5-enolpyruvyl-3-phosphoshikimic acid; herbicide; transgenic plant;
KW glyphosate resistance gene; glyphosate-tolerance; promoter.
XX
OS *Pseudomonas* sp. strain PG2982.
XX
FH Key Location/Qualifiers
FT Region 200..204
FT /label= characteristc_region
FT /note= "see AAW34690"
FT Region 26..29
FT /label= characteristc_region
FT /note= "see AAW34691"
FT Region 173..177
FT /label= characteristc_region
FT /note= "see AAW34692"
FT Region 273..276
FT /label= characteristc_region
FT /note= "see AAW34693"
XX
PN US5627061-A.
XX
XX 06-MAY-1997.

[illegible]

Db	361	vargleangvdcetegemsltvgrpdgkglg---ggtvathldhriamsflvmglaaekp	417
Qy	421	VTVDATMIATSFPEFMDLMAIGAKIELS	450
		: :	
Db	418	vtvdsmmiatsfpefmdmpglgakieIs	447
RESULT	13		
AAW24479			
ID	AAW24479	standard; Protein; 449 AA.	
XX	AC	AAW24479;	
XX	XX		
DT	02-OCT-1997	(first entry)	
XX	XX		
DE	Class II	EPSPS used for glyphosate resistant plant production.	
XX	KW	5-enolpyruvylshikimate-3-phosphate synthase; EPSPS; Class II;	
KW	KW	glyphosate resistant; transgenic plant; herbicide; shikimic acid;	
KW	KW	fusion protein; soybean; corn; wheat; barley; cotton; sugarbeet;	
XX	KW	tobacco.	
OS	Achromobacter	sp. strain LBAA.	
XX	XX		
FN	US5633435-A.		
XX	PD	27-MAY-1997.	
XX	XX		
PF	31-AUG-1990;	90US-0576537.	
XX	XX		
PR	13-SEP-1994;	94US-0306063.	
PR	31-AUG-1990;	90US-0576537.	
PR	28-AUG-1991;	91US-0749611.	
XX	XX		
PA	(MONS)	MONSANTO CO.	
XX	XX		
PI	Barry GF,	Kishore GM, Padgett SR, Stallings WC;	
XX	XX		
DR	WPI;	1997-297418/27.	
DR	N-PSDB;	AA77315.	
XX	XX		
PT	New isolated 5-enol:pyruvyl:shikimate-3-phosphate synthase gene -		
PT	used for transforming plants to produce plants which are tolerant to		
PT	glyphosate herbicide		
XX	XX		
PS	Claim 7;	Column 63-66; 154pp; English.	
XX	XX		
CC	AAW24479	shows the sequence of a class II 5-enolpyruvylshikimate-3-	
CC	phosphate synthase (EPSPS) enzyme. Class II EPSPS enzymes are tolerant		
CC	to glyphosate herbicides. EPSPS and sequences encoding it are used for		
CC	the production of herbicide resistant (glyphosate-tolerant) plants		
CC	such as corn, wheat, rice, barley, soybean, cotton, sugarbeet, oilseed		
CC	rape, canola, flax, sunflower, potato, tobacco, tomato, alfalfa, poplar,		
CC	pine, eucalyptus, apple, lettuce, peas, lentils, grape and turf grasses.		
XX	XX		
SQ	Sequence	449 AA;	
Query Match	83.38;	Score 1906.5;	DB 18; Length 449;
Best Local Similarity	83.18;	Pred. No. 1.1e-152;	
Matches	374;	Conservative	32; Mismatches 41; Indels 3; Gaps 1;
Qy	1	MSHGASSRPATKSSGSLGVNVRIPGDKSISHRFSMFGLASGETRITGLLEGEDVINTG	60
		: : : :	
Db	1	mshasapktacrrsealtgeiripgdkshrsfmfglasgetrlgllegedvintg	60
Qy	61	KAMQAMGARIRKEGDTWIIDGVNGGLLAPAPLDGNAATGCRLTMGLVGVDFTSTFI	120
		: : : : : : :	
Db	61	ramqamgakirkegdvwlingvgngcllqpeaaidfgnagtgarltmglygtymktsfi	120
Qy	121	GDASLTKRPNGRVLNPLREMGVQVKSEGDRLPVTILRGPKTPTPIYRVPMASQVKSAV	180
		: : : : : : :	
Db	121	qdaslskrpmrvinplremavgveaagddmptlligpktanplityrvpmasaqvksav	180

QY 181 LLAGLNTPGITTVIEPIIMTRDHTERKMLQGFANLTVETDADGVRTIRLEGRGKLTGQVID 240
 DB 181 lliagintpgyvtviepvmtrdhtekmlqgfadltvetdkdgvrrhritgqklygqtlid 240
 QY 241 VPGDPSSTAFPLVAALLVPGSDVTILNVLNPNTRTGLIILTLQEMGADIEVINPRLAGGED 300
 DB 241 vpgdpsstafplvaallvegsvdtrnvlmptrtglililqemgadievlnarlagged 300
 QY 301 VADLRVRSSTIKGVTVPEDRAPSMIDEPILAVAAFAEGATVMNGLLELRVKESDRLSA 360
 DB 301 vadlrvrasklkgvvppperapsmidelpvllaasfaegetvmdgldelrvkesdrlaa 360
 QY 361 VANGKLNGVDCDEGETSLVVRGPDGKGLGNASGAATVATHLDRHTAMSLVGLVSENP 420
 DB 361 vangleangvdcdegemsltvrrgpdgkglg---ggtvathldhriamslvmglvse 417
 QY 421 VTVDATMIATSPFPEFMDLMAGLAKIELS 450
 DB 418 vtvdsmnmiatsfpefmdmmpglgakiel 447
 RESULT 14
 AAW24480
 ID AAW24480 standard; Protein; 449 AA.
 AC AAW24480;
 DT 02-OCT-1997 (first entry)
 DE Class II EPSPS used for glyphosate resistant plant production.
 KW 5-enolpyruvylshikimate-3-phosphate synthase; EPSPS; Class II;
 KW glyphosate resistant; transgenic plant; herbicide; shikimic acid;
 KW fusion protein; soybean; corn; wheat; barley; cotton; sugarbeet;
 KW tobacco.
 OS Pseudomonas sp. strain PG2982.
 PN US5633435-A.
 PD 27-MAY-1997.
 PF 31-AUG-1990; 90US-0576537.
 PR 13-SEP-1994; 94US-0306063.
 PR 31-AUG-1990; 90US-0576537.
 PR 28-AUG-1991; 91US-0749611.
 PA (MONS) MONSANTO CO.
 PI Barry GF, Kishore GM, Padgett SR, Stallings WC;
 DR WPI; 1997-297418/27.
 DR N-PSDB; AAT77316.
 XX New isolated 5-enol:pyruvyl:shikimate-3-phosphate synthase gene -
 PT used for transforming plants to produce plants which are tolerant to
 PT glyphosate herbicide
 XX Claim 7; Column 69-72; 154pp; English.
 XX AAW24480 shows the sequence of a class II 5-enolpyruvylshikimate-3-
 CC phosphate synthase (EPSPS) enzyme. Class II EPSPS enzymes are tolerant
 CC to glyphosate herbicides. EPSPS and sequences encoding it are used for
 CC the production of herbicide resistant (glyphosate-tolerant) plants
 CC such as corn, wheat, rice, barley, soybean, cotton, sugarbeet, oilseed
 CC rape, canola, flax, sunflower, potato, tobacco, tomato, alfalfa, poplar,
 CC pine, eucalyptus, apple, lettuce, peas, lentils, grape and turf grasses.
 XX Sequence 449 AA;
 SQ

Query Match 83.3%; Score 1906.5; DB 18; Length 449;
 Best Local Similarity 83.1%; Pred. No. 1.1e-152;
 Matches 374; Conservative 32; Mismatches 41; Indels 3; Gaps 1;
 QY 1 MSHGASSRPATARKSSGLSGTVRIPODKSISHRSPWFGGLASGETRITGLLEGEDVING 60
 DB 1 mshsaspkpatarrsealtgeiripgdkisrsmfsglasgetritgllegedvinyg 60
 QY 61 KAMQAMGARIRREGDTWIIDGVNGGLLAPEAPLDFGNATGRLTMGLVGVYDFDSTFI 120
 DB 61 ramqamgarkirregdwiingvngcllpeaaldfgnagtgarlcmglvgytdmktfi 120
 QY 121 GDASLTKRPGRVLNPLREMGVQVKSDEGDRPLVTLRGPKTPTPTTYRVPMSAQVKSAV 180
 DB 121 gdaslskrpmgrvlnplremgvqveaadgrmpltligpktanpityrvpmsaqvksav 180
 QY 181 LLAGLNTPGITTVIEPIIMTRDHTERKMLQGFANLTVETDADGVRTIRLEGRGKLTGQVID 240
 DB 181 lliagintpgyvtviepvmtrdhtekmlqgfadltvetdkdgvrrhritgqklygqtlid 240
 QY 241 VPGDPSSTAFPLVAALLVPGSDVTILNVLNPNTRTGLIILTLQEMGADIEVINPRLAGGED 300
 DB 241 vpgdpsstafplvaallvegsvdtrnvlmptrtglililqemgadievlnarlagged 300
 QY 301 VADLRVRSSTIKGVTVPEDRAPSMIDEPILAVAAFAEGATVMNGLLELRVKESDRLSA 360
 DB 301 vadlrvrasklkgvvppperapsmidelpvllaasfaegetvmdgldelrvkesdrlaa 360
 QY 361 VANGKLNGVDCDEGETSLVVRGPDGKGLGNASGAATVATHLDRHTAMSLVGLVSENP 420
 DB 361 vangleangvdcdegemsltvrrgpdgkglg---ggtvathldhriamslvmglvse 417
 QY 421 VTVDATMIATSPFPEFMDLMAGLAKIELS 450
 DB 418 vtvdsmnmiatsfpefmdmmpglgakiel 447
 RESULT 15
 AAW71611
 ID AAW71611 standard; Protein; 449 AA.
 AC AAW71611;
 DT 14-DEC-1998 (first entry)
 DE Pseudomonas sp. strain PG2982 Class II EPSPS.
 KW 5-enolpyruvylshikimate-3-phosphate synthase; class II EPSPS enzyme;
 KW glyphosate herbicide; transformed bacteria; class I EPSPS enzyme;
 KW resistance; inhibition; 5-enolpyruvyl-3-phosphoshikimic acid;
 KW N-phosphonomethylglycine.
 OS Pseudomonas sp.
 PN US5804425-A.
 PD 08-SEP-1998.
 PF 07-APR-1997; 97US-0833485.
 PR 13-SEP-1994; 94US-0306063.
 PR 31-AUG-1990; 90US-0576537.
 PR 28-AUG-1991; 91US-0749611.
 PR 07-APR-1997; 97US-0833485.
 PA (MONS) MONSANTO CO.
 PI Barry GF, Kishore GM, Padgett SR, Stallings WC;
 DR WPI; 1998-505657/43.
 DR N-PSDB; AAV58012.
 XX Glyphosate resistant 5-enol:pyruvyl:shikimate-3-phosphate synthase -
 PT

PT useful for characterisation of the enzyme to determine inhibition
 PT data values

XX Disclosure: Fig 5; 152pp; English.

XX An isolated 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) has the
 CC sequence domains (I), (II), (III) and (IV): -R-X1-H-X2-E- (I),
 CC -G-D-K-X3- (II); -S-A-Q-X4-K- (III); and -N-X5-T-R- (IV). Where X1 = G,
 CC S, T, C, Y, N, Q, D or E; X2 and X3 = S or T; X4 and X5 = A, R, N, D, C,
 CC Q, E, G, H, I, L, K, M, F, P, S, T, W, Y or V. The EPSPS enzyme,
 CC produced by recombinant methods, can be used in kinetic studies to
 CC determine Ki and Km values of the enzyme for its characterisation. The
 CC enzyme is normally used for the production of 5-enolpyruvyl-3-
 CC phosphoshikimate acid in plants, and most forms of the enzyme are
 CC inhibited by N-phosphonomethylglycine ('glyphosate') herbicides.
 CC Inhibition data enables more accurate values of concentrations of
 CC herbicide to be used when growing the plant without being detrimental to
 CC it. This enables the plant to be grown in the presence of the herbicide,
 CC being used to inhibit the growth of undesired plants. The present
 CC sequence represents a Class II EPSPS from bacterial isolate
 CC Pseudomonas sp. strain PG2982.

XX Sequence 449 AA;

Query Match 83.3%; Score 1906.5; DB 19; Length 449;
 Best Local Similarity 83.1%; Pred. No. 1.1e-152;
 Matches 374; Conservative 32; Mismatches 41; Indels 3; Gaps 1;

QY 1 MSHGASSRPAATKSSGLSGTWRIPGDKISHSRPFMGGLASGETRITGLLEGEDVINTG 60
 DB 1 mshsaspkpatarsealtgeiripgdkisrfsfmgglasgetritgllegedvintg 60

QY 61 KAMQANGARIRKSGDWIWDGVNGGLLAPAPLDGNAATGCRMLGVGVYDFDSTFI 120
 DB 61 ranqangakirkegvdwiwngvngcllqpeaaldfgnagtarlmglvgytdmktffi 120

QY 121 GDASLTKRPMGRVNLPLREMGVGVKSGEDGRPLVTLRGPKTPITYRVPMSAQVKSAY 180
 DB 121 gdaslkrpmpgrvnlplremgvvgeadgdrmpitligpktpitpityrvpmsaqvksav 180

QY 181 LLAGLNTPGITTVIEPIIMTRDHTKMLQGFGANLTVETDADGVRTIRLEGRKLTQVID 240
 DB 181 llaglnpvgttviepvmtrdhtekmlqgfgadltvetdkdgvhrhritggklvgqtd 240

QY 241 VPGDPSSTAPPLVAALLVPGSDVTILNVLNPTRTGLILTQEMGADIEVNPRLAGGED 300
 DB 241 vpgdpsstaplvaallvegsvdtrnvlmnprrtgliltqemgadienvnrlagged 300

QY 301 VADLRVRSSTLKGVTVPEDRAPSMIDEPYPLAVAAFAEGATVMNGLELRKESDRLSA 360
 DB 301 vadlrvrssstlkgvtpedrapsmidepyplavaafaegatvmnglelrkesdrlsa 360

QY 361 VANGKLNGVDCDEGETSLVVRGRPGCKGLGNASGAATVATLHDRIAMSLVMGLVSENP 420
 DB 361 vangleangvdcdegetslvvrgrpgckglgnasgaatvatldhriamslvmglvse 420

QY 421 VTVDATMIATSPPEFMDLMAGLAKIELS 450
 DB 418 vtvdsmiatstspfmmdmmpglgakiel 447

Search completed: August 15, 2002, 13:58:47
 Job time: 66 sec

THIS PAGE BLANK (USPTO)